

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 02-04-03
Searcher: Beverly E4994
Terminal time: 20
Elapsed time: _____
CPU time: _____
Total time: 25
Number of Searches: _____
Number of Databases: 1

Search Site

_____ STIC
_____ CM-1
_____ Pre-S

Type of Search

_____ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

Vendors

_____ IG Suite
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other CGN

THIS PAGE BLANK (USPTO)

PT activity than insulin -
 XX
 PS Example 1; Page 7; 24pp; English.
 CC The invention relates to a single-chain insulin analogue (SIA) compound
 CC comprising the properties of greater insulin receptor binding activity
 CC than proinsulin and less insulin receptor binding activity than insulin.
 CC The SIA compound has the formula: B chain - X - A chain; B and A = human
 CC insulin chains or functional analogues; and X = a joining peptide of 5 to
 CC 18 amino acids. The SIA compound or a vector comprising a polynucleotide
 CC encoding the SIA compound is used to treat type I diabetes mellitus. The
 CC present sequence represents a linker region of a SIA-1 sequence.

XX SQ Sequence 7 AA;

Query Match 100.0%; Score 41; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
 Db 1 GGGPGKR 7

RESULT 2

AAR90538
 ID AAR90538 standard; protein; 135 AA.

XX AC AAR90538;

XX DT 08-AUG-1996 (first entry)

XX DE pJG4-5-CDK-BP clone #68 derived CDK4 binding protein.

XX KW Cell cycle; CDK4; regulation; G1 phase; proliferation; tumourigenesis;
 KW cyclin dependent kinase; differentiation; CDK4 inhibitor; agonist;
 KW antagonist.

XX OS Synthetic.

XX PN W09533819-A2.

XX PD 14-DEC-1995.

XX PF 02-JUN-1995; 95WO-US07113.

XX PR 02-JUN-1994; 94US-0253155.

XX PA (MITO-) MITOTIX INC.

XX PI Draetta G, Gyuris J;

XX DR WPI: 1996-040227/04.

XX DR N-PSDB; AAT12167.

PT Cyclin-dependent kinase-4 binding protein - used in the isolation of
 PT (ant)agonists of cell cycle regulation.

PS Claim 1; Page 73; 115pp; English.

XX AAR90533-R90556 are cyclin dependent kinase 4 (CDK4) binding proteins.
 CC encoded by clones of the plasmid pJG4-5-CDKBP. CDK4 binding proteins
 CC (CDK4-BP) may be used in an assay for screening test compounds as
 CC inhibitors of CDK/CDK4-BP interaction. The complexes formed by CDK4
 CC and D-type cyclins are strongly implicated in the control of the early
 CC G1 phase of the cell cycle and are strong candidates for controlling
 CC and/or preventing tumourigenesis and the onset of cancer. Nucleic acids
 CC encoding CDK4-BP or fragments of these may be used as probes/primers
 CC to diagnose the presence or absence of genetic lesions in a gene
 CC encoding 1 of the 24 CDK4-BP, and hence to diagnose the risk for a
 CC subject of developing a cell-proliferation associated disorder (e.g.
 CC cancer).

XX

SQ Sequence 135 AA;

Query Match 92.7%; Score 38; DB 17; Length 135;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
 Db 82 GGGPGRR 88

RESULT 3

AAB53475

ID AAB53475 standard; Protein; 573 AA.

XX AC AAB53475;

XX DT 09-MAR-2001 (first entry)

XX DE Human colon cancer antigen protein sequence SEQ ID NO:1015.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;
 KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
 KW neural disorder; immune system disorder; muscular disorder;
 KW reproductive disorder; gastrointestinal disorder; renal disorder;
 KW infectious disease; cardiovascular disorder.

XX OS Homo sapiens.

XX PN W0200055351-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05883.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI: 2000-587534/55.

XX DR N-PSDB; AAC98232.

PT Colon cancer associated gene sequences, referred to as colon cancer
 PT antigens, useful for the treatment, prevention, and diagnosis of colon
 PT disorders such as colon cancer -

PS Claim 11; Page 1600-1602; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
 CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
 CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
 CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,
 CC proteins and antibodies to the proteins are useful for the prevention,
 CC treatment and diagnosis of colon disorders, such as colon cancer. The
 CC polynucleotides may be used in diagnostics and research, such as for
 CC chromosome identification, and as hybridisation probes. The proteins
 CC may also be used to prevent diseases such as neural disorders, immune
 CC system disorders, muscular disorders, wounds, renal disorders, infectious
 CC gastrointestinal disorders, reproductive disorders,
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
 CC AAB54007 represent sequences used in the exemplification of the present
 CC invention.

XX SQ Sequence 573 AA;

Query Match 92.7%; Score 38; DB 21; Length 573;
 Best Local Similarity 85.7%; Pred. No. 4e+02;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 4, 2003, 07:01:48 ; Search time 69 Seconds
(without alignments)
13.518 Million cell updates/sec

Title: US-09-706-690-1
Perfect score: 41
Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 41 | 100.0 | 7 | 23 | Single-chain insul |
| 2 | 38 | 92.7 | 135 | 17 | pJG4-5-CDK-Bp clon |
| 3 | 38 | 92.7 | 573 | 21 | Human colon cancer |
| 4 | 38 | 92.7 | 633 | 22 | C glutamicum prote |
| 5 | 38 | 92.7 | 782 | 22 | Human polypeptide |
| 6 | 38 | 92.7 | 886 | 22 | Human protein sequ |
| 7 | 38 | 92.7 | 972 | 22 | Hamster EST encode |
| 8 | 38 | 92.7 | 972 | 22 | Hamster EST encode |
| 9 | 36 | 87.8 | 39 | 22 | Human nervous syst |
| 10 | 36 | 87.8 | 115 | 22 | Novel human secret |

| | | | | | | |
|----|----|------|-----|----|----------|--------------------|
| 11 | 36 | 87.8 | 119 | 22 | AA674224 | Human colon cancer |
| 12 | 36 | 87.8 | 132 | 22 | AA051115 | Human polypeptide |
| 13 | 36 | 87.8 | 154 | 22 | AAU48315 | Propionibacterium |
| 14 | 36 | 87.8 | 158 | 22 | AA693255 | Human protein HP10 |
| 15 | 36 | 87.8 | 533 | 22 | AAE05669 | Triglochin maritim |
| 16 | 36 | 87.8 | 540 | 22 | AAE05668 | Triglochin maritim |
| 17 | 35 | 85.4 | 113 | 22 | AA02405 | Human polypeptide |
| 18 | 35 | 85.4 | 153 | 20 | AAV74188 | Human prostate tum |
| 19 | 35 | 85.4 | 179 | 22 | AAW25820 | Human protein sequ |
| 20 | 35 | 85.4 | 296 | 21 | AA842246 | Human ORFX ORF2010 |
| 21 | 35 | 85.4 | 320 | 22 | AA692504 | C glutamicum prote |
| 22 | 35 | 85.4 | 387 | 10 | AA693672 | Sequence from acyB |
| 23 | 35 | 85.4 | 485 | 21 | AA659032 | Breast and ovarian |
| 24 | 35 | 85.4 | 591 | 20 | AAV15228 | Human receptor pro |
| 25 | 35 | 85.4 | 594 | 19 | AAW72094 | HSV-2 strain SB5 C |
| 26 | 35 | 85.4 | 602 | 19 | AAW72001 | HSV-2 strain SB5 C |
| 27 | 35 | 85.4 | 713 | 20 | AAV41712 | Human PRO724 prote |
| 28 | 35 | 85.4 | 713 | 21 | AA844268 | Human PRO724 (UNQ3 |
| 29 | 35 | 85.4 | 713 | 21 | AAV71081 | Human TANGO 136 pr |
| 30 | 35 | 85.4 | 713 | 22 | AAU29231 | Human PRO polypept |
| 31 | 35 | 85.4 | 713 | 23 | AB895462 | Human angiogenesis |
| 32 | 35 | 85.4 | 713 | 23 | AB890346 | Human polypeptide |
| 33 | 35 | 85.4 | 713 | 23 | AB884856 | Human PRO724 prote |
| 34 | 35 | 85.4 | 713 | 23 | AB805751 | Human G protein-co |
| 35 | 35 | 85.4 | 718 | 19 | AAW72245 | HSV-2 strain SB5 C |
| 36 | 35 | 85.4 | 728 | 19 | AAW72246 | HSV-2 strain SB5 C |
| 37 | 35 | 85.4 | 771 | 19 | AAW72247 | HSV-2 strain SB5 C |
| 38 | 35 | 85.4 | 805 | 19 | AAW72248 | HSV-2 strain SB5 C |
| 39 | 35 | 85.4 | 818 | 19 | AAW72249 | HSV-2 strain SB5 C |
| 40 | 35 | 85.4 | 821 | 19 | AAW72250 | HSV-2 strain SB5 C |
| 41 | 34 | 82.9 | 91 | 22 | AB867032 | Drosophila melanog |
| 42 | 34 | 82.9 | 132 | 22 | AA007877 | Human polypeptide |
| 43 | 34 | 82.9 | 195 | 23 | ABG59989 | Human DITHP polype |
| 44 | 34 | 82.9 | 258 | 22 | ABG11350 | Novel human diagno |
| 45 | 34 | 82.9 | 354 | 22 | ABG23888 | Novel human diagno |

ALIGNMENTS

RESULT 1
ABB80744
ID ABB80744 standard; peptide; 7 AA.
AC ABB80744;
XX
XX
DT 15-JUL-2002 (first entry)
XX
DE Single-chain insulin analogue (SIA)-1 linker sequence.
XX
KW Single-chain insulin analogue; SIA; insulin; proinsulin; antidiabetic;
KW SIA-1; gene therapy; linker.
XX
OS Synthetic.
XX
PN EPI193272-A1.
XX
PD 03-APR-2002.
XX
PF 13-SEP-2001; 2001EP-0121651.
XX
PR 02-OCT-2000; 2000KR-0058003.
PR 07-NOV-2000; 2000US-0706690.
(LEE/H/) LEE H C.
Lee HC, Kim S, Kim K, Shin H, Yoon J;
WPI; 2002-373742/41.
A single chain insulin analog (SIA) compound used in the treatment of type I diabetes, comprises the properties of greater insulin receptor binding activity than proinsulin and less insulin receptor binding

```

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-0116126.
XX PR 29-JUL-1999; 99JP-0248036.
XX PR 27-AUG-1999; 99JP-0300253.
XX PR 11-JAN-2000; 2000JP-0118776.
XX PR 02-MAY-2000; 2000JP-0183767.
XX PR 09-JUN-2000; 2000JP-0241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX PR Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX Claim 8; SEQ ID 14101; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX SQ Sequence 886 AA;

Query Match 92.7%; Score 38; DB 22; Length 886;
Best Local Similarity 85.7%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||
Db 350 GGGPGR 356

RESULT 7
AAM23988
ID AAM23988 standard; Protein; 972 AA.
XX AC AAM23988;
XX 12-OCT-2001 (first entry)
XX
XX

```

```

DE XX Hamster EST encoded protein SEQ ID NO: 1513.
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition.
XX OS Cricetulus griseus.
XX PN WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US02687.
XX PR 25-JAN-2000; 2000US-0491404.
XX PR 17-JUL-2000; 2000US-0617746.
XX PR 03-AUG-2000; 2000US-0631451.
XX PR 15-SEP-2000; 2000US-0663870.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX WPI; 2001-476164/51.
XX DR N-PSDB; AAH98647.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use -
XX PS Claim 20; Page 1043-1045; 1275pp; English.
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
XX protein of the invention.
XX SQ Sequence 972 AA;

Query Match 92.7%; Score 38; DB 22; Length 972;
Best Local Similarity 85.7%; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||
Db 435 GGGPGR 441

RESULT 8
AAM24010
ID AAM24010 standard; Protein; 972 AA.
XX AC AAM24010;
XX 12-OCT-2001 (first entry)
XX
XX Hamster EST encoded protein SEQ ID NO: 1535.
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder;
KW biodiversity; gene therapy; nutrition.
XX OS Cricetulus griseus.
XX PN WO200154477-A2.
XX PD 02-AUG-2001.

```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||:1

Db 36 GGGPGRR 42

RESULT 4

AAG92232
ID AAG92232 standard; Protein; 633 AA.
XX
AC AAG92232;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 5986.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
DR N-PSDB; AAH67451.

XX
PT Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 5986; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.

SQ Sequence 633 AA;

Query Match 92.7%; Score 38; DB 22; Length 633;
Best Local Similarity 85.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||:1

Db 105 GGGPGRR 111

RESULT 5

AAB93996
ID AAB93996 standard; Protein; 886 AA.
XX
AC AAB93996;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:14101.

AAO00740
ID AAO00740 standard; Protein; 782 AA.
XX
AC AAO00740;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 14632.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
DR N-PSDB; AAI80671.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 14632; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 782 AA;

Query Match 92.7%; Score 38; DB 22; Length 782;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||:1

Db 374 GGGPGRR 380

PR 20-OCT-2000; 2000US-0240950.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-02559678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-541565/60.
DR N-PSDB; ABA11905.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Claim 11; SEQ ID NO 4236; 1701pp + Sequence Listing; English.
PS
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABB14678-ABB18001) useful for preventing, treating or ameliorating
*CC

CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 39 AA;

Query Match 87.8%; Score 36; DB 22; Length 39;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGK 6
Db 29 GGGPGK 34
|||||

RESULT 10
AAU32518
ID AAU32518 standard; Protein; 115 AA.
XX
XX AAU32518;
XX
XX 18-DEC-2001 (first entry)
XX
XX Novel human secreted protein #3009.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
OS Homo sapiens.
XX
XX WO200179449-A2.
XX
XX 25-OCT-2001.
XX
XX 16-APR-2001; 2001WO-US08656.
XX
XX 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-611725/70.
XX
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
PT
XX Claim 20; Page 626; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically

```
XX 25-JAN-2001; 2001WO-US02687.
PF 25-JAN-2000; 2000US-0491404.
XX 17-JUL-2000; 2000US-0617746.
PR 03-AUG-2000; 2000US-0631451.
PR 15-SEP-2000; 2000US-0663870.
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX WPI; 2001-476164/51.
DR N-PSDB; AAH98669.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use -
XX
XX Claim 20; Page 1056-1059; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.
XX
XX Sequence 972 AA;
SQ
Query Match 92.78; Score 38; DB 22; Length 972;
Best Local Similarity 85.78; Pred. No. 6.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGKGR 7
|||||
Db 435 GGGGRR 441
RESULT 9
ABBI5579
ID ABB15579 standard; Protein; 39 AA.
XX
AC ABB15579;
XX
XX 23-JAN-2002 (first entry)
XX
XX Human nervous system related polypeptide SEQ ID NO 4236.
XX
XX Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antidiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229267.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
```


QY 1 GGGPGK 6
|
Db 68 GGGPGK 73

RESULT 13
AAU48315
ID AAU48315 standard; Protein; 154 AA.
XX

AC AAU48315;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #9211.

XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypervitosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX
OS Propionibacterium acnes.

XX
PN WO200181581-A2.

XX
PD 01-NOV-2001.

XX
PF 20-APR-2001; 2001WO-US12865.

XX
PR 21-APR-2000; 2000US-199047P.

XX
PR 02-JUN-2000; 2000US-208841P.

XX
PR 07-JUL-2000; 2000US-216747P.

XX
PA (CORI-) CORIXA CORP.

XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX
DR WPI; 2001-616774/71.

XX
DR N-PSDB; AAS9542.

XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX
PS Example 1; SEQ ID No 9510; 1069pp; English.

XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypervitosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).

XX
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 154 AA;

Query Match 87.8%; Score 36; DB 22; Length 154;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6
|
Db 37 GGGPGK 42

RESULT 14
AAG93255
ID AAG93255 standard; Protein; 158 AA.
XX

AC AAG93255;

DT 13-SEP-2001 (first entry)

DE Human protein HP10321.

XX
KW Human; gene therapy; tumour.

XX
OS Homo sapiens.

XX
PN WO200142302-A1.

XX
PD 14-JUN-2001.

XX
PF 06-DEC-2000; 2000WO-JP08631.

XX
PR 06-DEC-1999; 99JP-0346863.

XX
PR 06-DEC-1999; 99JP-0346864.

XX
PR 08-FEB-2000; 2000JP-0031062.

XX
PR 10-FEB-2000; 2000JP-0034090.

XX
PR 10-FEB-2000; 2000JP-0034091.

XX
PR 14-FEB-2000; 2000JP-0035829.

XX
PR 14-FEB-2000; 2000JP-0071161.

XX
PR 30-MAY-2000; 2000JP-0160851.

XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX
PI Kato S, Eguchi C, Saeki M;

XX
PI WPI; 2001-381646/40.

XX
DR N-PSDB; AAH68540.

XX
PT Human protein originated from tumor cell line, applicable as drug,
PT reagent for studying intracellular protein networks and protein source
PT for drug screening, also encoded cDNA for gene diagnosis and gene
PT therapy -

XX
PS Claim 1; Pages 133-134; 471pp; Japanese.

XX
CC The present sequence is a human protein. The human protein, preferably
CC originated from tumour cell line, is applicable as a drug, a reagent for
CC studying intracellular protein networks and a protein source for
CC screening proteins for binding low molecular weight drugs. The human
CC protein coding sequence is useful for gene diagnosis and gene therapy,
CC expression vectors and transformant cells for detection of ligands and
CC receptors.

XX
SQ Sequence 158 AA;

Query Match 87.8%; Score 36; DB 22; Length 158;

Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
AAE05669
ID AAE05669 standard; Protein; 533 AA.
XX

CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX Sequence 115 AA;

Query Match 87.8%; Score 36; DB 22; Length 115;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGK 6
Db 75 GGGPGK 80
|||||

RESULT 11
AAG74224

ID AAG74224 standard; Protein: 119 AA.

XX AAG74224;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:4988.

XX Human: colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.

XX Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI; 2001-235357/24.

DR N-PSDB; AAH33655.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -

PS Claim 11; Page 6735-6736; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH77789 represent sequences used in the exemplification of the
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were

CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 119 AA;

Query Match 87.8%; Score 36; DB 22; Length 119;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGK 6
Db 90 GGGPGK 95
|||||

RESULT 12

AAO05115

ID AAO05115 standard; Protein: 132 AA.

XX AAO05115;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 19007.

XX Human: cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

PN WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR N-PSDB; AA185046.

XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 19007; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 132 AA;

Query Match 87.8%; Score 36; DB 22; Length 132;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

RESULT 3
PCT-US91-07635-6
; Sequence 6, Application PC/TUS9107635
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: RUEGER, DAVID C
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/07635
; FILING DATE: 19911018
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-0566PC
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US91-07635-6

Query Match 82.9%; Score 34; DB 5; Length 399;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

RESULT 4
US-07-841-646-29
; Sequence 29, Application US/07841646
; Patent No. 5266683
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,646
; FILING DATE: 19920221
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-841-646-29

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 09:28:13 ; Search time 23 Seconds
(without alignments)
8.955 Million cell updates/sec

Title: US-09-706-690-1

Perfect score: 41

Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|-------|---------|--------------|-------|----------------------|
| 1 | 35 | 85.4 | 387 | 6 | Patent No. 5240849-5 |
| 2 | 35 | 85.4 | 714 | 4 | US-09-347-878-16 |
| 3 | 34 | 82.9 | 399 | 5 | PCT-US91-07635-6 |
| 4 | 34 | 82.9 | 402 | 1 | US-07-841-646-29 |
| 5 | 34 | 82.9 | 402 | 1 | US-07-901-703-11 |
| 6 | 34 | 82.9 | 402 | 1 | US-08-147-023-29 |
| 7 | 34 | 82.9 | 402 | 1 | US-08-206-864-4 |
| 8 | 34 | 82.9 | 402 | 1 | US-08-278-729A-21 |
| 9 | 34 | 82.9 | 402 | 1 | US-08-480-528A-8 |
| 10 | 34 | 82.9 | 402 | 1 | US-08-479-666-8 |
| 11 | 34 | 82.9 | 402 | 1 | US-08-155-343A-21 |
| 12 | 34 | 82.9 | 402 | 1 | US-08-406-672-21 |
| 13 | 34 | 82.9 | 402 | 1 | US-08-643-563A-21 |
| 14 | 34 | 82.9 | 402 | 1 | US-08-447-570-29 |
| 15 | 34 | 82.9 | 402 | 1 | US-08-643-763A-21 |
| 16 | 34 | 82.9 | 402 | 1 | US-08-462-623-21 |
| 17 | 34 | 82.9 | 402 | 1 | US-08-451-953A-21 |
| 18 | 34 | 82.9 | 402 | 2 | US-08-459-346-6 |
| 19 | 34 | 82.9 | 402 | 2 | US-08-445-468A-21 |
| 20 | 34 | 82.9 | 402 | 2 | US-08-901-200A-8 |
| 21 | 34 | 82.9 | 402 | 2 | US-08-449-700-29 |
| 22 | 34 | 82.9 | 402 | 2 | US-08-449-699A-29 |
| 23 | 34 | 82.9 | 402 | 2 | US-08-461-397A-21 |
| 24 | 34 | 82.9 | 402 | 2 | US-08-912-088-21 |
| 25 | 34 | 82.9 | 402 | 2 | US-08-278-730A-21 |
| 26 | 34 | 82.9 | 402 | 3 | US-08-889-419-6 |
| 27 | 34 | 82.9 | 402 | 3 | US-08-445-467-21 |

```

28 34 82.9 402 3 US-08-480-515A-21 Sequence 21, Appl
29 34 82.9 402 4 US-09-219-391-8 Sequence 8, Appl
30 34 82.9 402 4 US-09-170-336-21 Sequence 21, Appl
31 34 82.9 402 4 US-08-402-542-6 Sequence 6, Appl
32 34 82.9 402 4 US-08-461-113-21 Sequence 21, Appl
33 34 82.9 402 5 PCT-US92-01968-21 Sequence 11, Appl
34 34 82.9 402 5 PCT-US93-05446-11 Sequence 6, Appl
35 34 82.9 402 5 PCT-US93-07189-6 Sequence 21, Appl
36 34 82.9 402 5 PCT-US93-07190-21 Sequence 21, Appl
37 34 82.9 402 5 PCT-US93-07231-21 Sequence 21, Appl
38 34 82.9 402 5 PCT-US93-08742-21 Sequence 21, Appl
39 34 82.9 402 5 PCT-US93-08808-21 Sequence 21, Appl
40 34 82.9 402 5 PCT-US93-08885-21 Sequence 21, Appl
41 34 82.9 402 5 PCT-US93-10520-8 Sequence 8, Appl
42 34 82.9 830 4 US-09-562-737-37 Sequence 37, Appl
43 33 80.5 87 4 US-09-134-001C-3031 Sequence 3031, Ap
44 33 80.5 154 4 US-09-723-830-2 Sequence 2, Appl
45 33 80.5 426 3 US-08-826-246-4 Sequence 4, Appl

```

ALIGNMENTS

```

RESULT 1
5240849-5
; Patent No. 5240849
; APPLICANT: Arisawa, Akira; Kawamura, Naoto; Kojima, Ikou; Okumura,
; Yasushi; Kazuhiko, Okamura; Hiroshi, Tone; Okamura Rokuro
; TITLE OF INVENTION: DNA CODING FOR ENZYME CAPABLE OF ACILYATING
; THE "A"-POSITION OF MACROLIDE ANTIBIOTIC
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/06/353,323
; FILING DATE: 23-MAY-1989
; SEQ ID NO: 5
; LENGTH: 387
5240849-5

```

Query Match 85.4%; Score 35; DB 6; Length 387;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 GGGPGKR 7
Db 165 GGGPGDR 171

```

```

RESULT 2
US-09-347-878-16
; Sequence 16, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 714
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-347-878-16

```

Query Match 85.4%; Score 35; DB 4; Length 714;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GGGPGKR 7
Db 46 GGGPGRK 52

```

Db 20 GGGPGLR 26
||||| I
RESULT 7
US-08-206-864-4
; Sequence 4, Application US/08206864
; Patent No. 5610021
; GENERAL INFORMATION:
; APPLICANT: RUEGER, DAVID C
; APPLICANT: JONES, WILLIAM K
; APPLICANT: TUCKER, RONALD F
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; RECOMBINANT OSTEOGENIC PROTEIN PRODUCTION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
; ADDRESSEE: INC.
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/206,864
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/027,070
; FILING DATE: 04-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/841,646
; FILING DATE: 21-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-096
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-206-864-4
Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGPKR 7
Db 20 GGGPGLR 26
||||| I
RESULT 8
US-08-278-729A-21
; Sequence 21, Application US/08278729A
; Patent No. 5650276
; GENERAL INFORMATION:
; APPLICANT: SMART, JOHN
; APPLICANT: OPPERMAN, HERMAN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,528A
; FILING DATE: 20-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER Esq., EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-058CPFW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-278-729A-21
Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGPKR 7
Db 20 GGGPGLR 26
||||| I
RESULT 9
US-08-480-528A-8
; Sequence 8, Application US/08480528A
; Patent No. 5652118
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,528A

RESULT 5
US-07-901-703-11
; Sequence 11, Application US/07901703
; Patent No. 5344654
; GENERAL INFORMATION:
; APPLICANT: RUEGER, DAVID C
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZAKYNAK, ENGIN
; TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: EXCHANGE PLACE, 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/901.703
; FILING DATE: 19920616
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, ESQ, EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: STK-057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-901-703-11
Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 GGGPGKR 7
Db 20 GGGPGLR 26
IIIIII
RESULT 6
US-08-147-023-29
; Sequence 29, Application US/08147023
; Patent No. 5468845
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/147,023
; FILING DATE: 21-FEB-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-147-023-29
Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 GGGPGKR 7


```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/406,672
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 752,857
; FILING DATE: 30-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 667,274
; FILING DATE: 11-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-060CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-406-672-21

Query Match      82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGGPGKR 7
Db      20 GGGPGLR 26

RESULT 13
US-08-643-563A-21
; Sequence 21, Application US/08643563A
; Patent No. 5707810
; GENERAL INFORMATION:
; APPLICANT: SMART, JOHN
; APPLICANT: OPPERMANN, HERMAN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US/08/643,563A
; FILING DATE: 06-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TWOMEY ESQ., MICHAEL J.
; REGISTRATION NUMBER: 38,349
; REFERENCE/DOCKET NUMBER: CRP-058CN2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-643-563A-21

Query Match      82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GGGPGKR 7
Db      20 GGGPGLR 26

RESULT 14
US-08-447-570-29
; Sequence 29, Application US/08447570
; Patent No. 5714589
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,570
; FILING DATE: 21-FEB-1992
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
```

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-076FW
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-480-528A-8

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

RESULT 10
US-08-479-666-8
; Sequence 8, Application US/08479666
; Patent No. 5652337
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; APPLICANT: COHEN, CHARLES M.
; TITLE OF INVENTION: OP3-INDUCED MORPHOGENESIS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/479,666
; APPLICATION NUMBER: US/08/479,666
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-076DV
; TELEPHONE: (508) 435-9001
; TELEFAX: (508) 435-6951
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-479-666-8

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

RESULT 11
US-08-155-343A-21
; Sequence 21, Application US/08155343A
; Patent No. 5656593
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: OPPERMAN, HERMAN
; APPLICANT: COHEN, CHARLES M.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE
; REGENERATION.
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/155,343A
; APPLICATION NUMBER: US/08/155,343A
; FILING DATE: 15-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-067FW
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-155-343A-21

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

RESULT 12
US-08-406-672-21
; Sequence 21, Application US/08406672
; Patent No. 5674844
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: COHEN, CHARLES M.
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: TREATMENT TO PREVENT LOSS OF AND/OR
; INCREASE BONE MASS IN METABOLIC BONE DISEASES
; NUMBER OF SEQUENCES: 33

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-570-29

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

RESULT 15
US-08-643-763A-21
; Sequence 21, Application US/08643763A
; Patent No. 5733878
; GENERAL INFORMATION:
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: OPPERMANN, HERMAN
; APPLICANT: COHEN, CHARLES M.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 01748
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/643,763A
; FILING DATE: 06-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTON ESQ., GILLIAN M.
; REGISTRATION NUMBER: 36,508
; REFERENCE/DOCKET NUMBER: CRP-067CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7560
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-643-763A-21

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 20 GGGPGLR 26

Search completed: February 4, 2003, 09:33:10
Job time : 26 secs

THIS PAGE BLANK (USPTO)

US-09-738-626-5986
; Sequence 5986, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIKO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5986
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5986

Query Match 92.7%; Score 38; DB 9; Length 633;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
DB 105 GGGPGR 111

RESULT 3
US-09-950-370-22
; Sequence 22, Application US/09950370
; Patent No. US20020076764A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Meyers, Rachel
; APPLICANT: Silos-Santiago, Inmaculad
; TITLE OF INVENTION: 27877, A No. US20020076764A1 Human Phospholipase and Uses There
; FILE REFERENCE: 10147-43U1
; CURRENT APPLICATION NUMBER: US/09/950,370
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,084
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-950-370-22

Query Match 90.2%; Score 37; DB 10; Length 875;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
DB 137 GGGPGR 143

RESULT 4
US-10-067-813-17
; Sequence 17, Application US/10067813
; Patent No. US20020156013A1
; GENERAL INFORMATION:
; APPLICANT: Renauld, Jean-Christophe
; APPLICANT: Louahed, Jamila
; APPLICANT: Grasso, Luigi
; APPLICANT: Levitt, Roy
; APPLICANT: Nicolaides, Nicholas
; TITLE OF INVENTION: Asthma Associated Factors as Targets for Treating
; FILE REFERENCE: 036870-5071
; CURRENT APPLICATION NUMBER: US/10/067,813
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/157,247
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 218
; TYPE: PRT
; ORGANISM: R-Ras
US-10-067-813-17

Query Match 85.4%; Score 35; DB 9; Length 218;
Best Local Similarity 85.7%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
DB 16 GGGPGR 22

RESULT 5
US-09-738-626-6258
; Sequence 6258, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIKO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6258
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6258

Query Match 85.4%; Score 35; DB 9; Length 320;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 09:29:37 ; Search time 18 seconds
(without alignments)
8.621 Million cell updates/sec

Title: US-09-706-690-1

Perfect score: 41

Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|--------------------|
| 1 | 38 | 92.7 | 573 | 10 | US-09-925-299-1015 |
| 2 | 38 | 92.7 | 633 | 9 | Sequence 1015, Ap |
| 3 | 37 | 90.2 | 875 | 10 | Sequence 5986, Ap |
| 4 | 35 | 85.4 | 218 | 9 | Sequence 22, Appl |
| 5 | 35 | 85.4 | 320 | 9 | Sequence 17, Appl |
| 6 | 35 | 85.4 | 713 | 9 | Sequence 6258, Ap |
| 7 | 35 | 85.4 | 713 | 9 | Sequence 183, App |
| 8 | 35 | 85.4 | 713 | 9 | Sequence 183, App |
| 9 | 35 | 85.4 | 713 | 9 | Sequence 183, App |
| 10 | 35 | 85.4 | 713 | 9 | Sequence 183, App |
| 11 | 35 | 85.4 | 713 | 9 | Sequence 183, App |
| 12 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 13 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 14 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 15 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 16 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 17 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 18 | 35 | 85.4 | 713 | 9 | Sequence 416, App |
| 19 | 35 | 85.4 | 713 | 9 | Sequence 416, App |

| | | | | | | |
|----|----|------|------|----|---------------------|-------------------|
| 20 | 35 | 85.4 | 713 | 9 | US-10-180-552-416 | Sequence 416, App |
| 21 | 35 | 85.4 | 713 | 9 | US-10-180-557-416 | Sequence 416, App |
| 22 | 35 | 85.4 | 713 | 12 | US-10-052-586-416 | Sequence 416, App |
| 23 | 34 | 82.9 | 378 | 10 | US-09-764-864-1240 | Sequence 1240, Ap |
| 24 | 34 | 82.9 | 403 | 9 | US-09-813-398-31 | Sequence 31, Appl |
| 25 | 34 | 82.9 | 861 | 10 | US-09-804-551B-20 | Sequence 4, Appl |
| 26 | 33 | 80.5 | 426 | 8 | US-08-825-486-4 | Sequence 8, Appl |
| 27 | 33 | 80.5 | 426 | 8 | US-08-870-434-8 | Sequence 4, Appl |
| 28 | 33 | 80.5 | 426 | 10 | US-09-372-044-4 | Sequence 4, Appl |
| 29 | 33 | 80.5 | 426 | 10 | US-09-524-417-67 | Sequence 67, Appl |
| 30 | 33 | 80.5 | 626 | 10 | US-09-862-027-43 | Sequence 43, Appl |
| 31 | 33 | 80.5 | 627 | 10 | US-09-862-027-46 | Sequence 46, Appl |
| 32 | 33 | 80.5 | 627 | 10 | US-09-862-027-47 | Sequence 47, Appl |
| 33 | 33 | 80.5 | 628 | 10 | US-09-862-027-48 | Sequence 48, Appl |
| 34 | 32 | 78.0 | 41 | 10 | US-09-864-761-38559 | Sequence 38559, A |
| 35 | 32 | 78.0 | 188 | 9 | US-09-989-920-240 | Sequence 240, App |
| 36 | 32 | 78.0 | 340 | 10 | US-09-900-237-22 | Sequence 22, Appl |
| 37 | 32 | 78.0 | 399 | 10 | US-09-764-864-891 | Sequence 891, App |
| 38 | 32 | 78.0 | 404 | 10 | US-09-191-687B-2 | Sequence 2, Appl |
| 39 | 32 | 78.0 | 441 | 9 | US-09-738-626-5934 | Sequence 5934, Ap |
| 40 | 32 | 78.0 | 468 | 9 | US-09-975-719-29 | Sequence 29, Appl |
| 41 | 32 | 78.0 | 530 | 9 | US-09-738-626-4780 | Sequence 4780, Ap |
| 42 | 32 | 78.0 | 855 | 10 | US-09-817-913-13 | Sequence 13, Appl |
| 43 | 32 | 78.0 | 855 | 10 | US-09-817-538-13 | Sequence 13, Appl |
| 44 | 32 | 78.0 | 1165 | 10 | US-09-900-237-8 | Sequence 8, Appl |
| 45 | 31 | 75.6 | 5 | 10 | US-09-269-439-9 | Sequence 9, Appl |

ALIGNMENTS

RESULT 1

US-09-925-299-1015

; Sequence 1015, Application US/09925299

; Patent No. US20020055627A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1015

; LENGTH: 573

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (14)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (28)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE

; LOCATION: (179)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-299-1015

Query Match 92.7%; Score 38; DB 10; Length 573;
Best Local Similarity 85.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | | : |
Db 36 GGGPGRR 42

RESULT 2

Db 53 GGPGR 59

RESULT 6

US-09-978-295A-183

; Sequence 183, Application US/09978295A

; Patent No. US20020156006A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kijavlin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630P1C11

; CURRENT APPLICATION NUMBER: US/09/978,295A

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077641

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077791

; PRIOR FILING DATE: 1998-03-12

; PRIOR APPLICATION NUMBER: 60/078004

; PRIOR FILING DATE: 1998-03-13

; PRIOR APPLICATION NUMBER: 60/078886

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078936

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078939

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/079294

; PRIOR FILING DATE: 1998-03-25

; PRIOR APPLICATION NUMBER: 60/079656

; PRIOR FILING DATE: 1998-03-26

; PRIOR APPLICATION NUMBER: 60/079664

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079689

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079663

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079728

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079786

; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079920

; PRIOR FILING DATE: 1998-03-30

; PRIOR APPLICATION NUMBER: 60/079923

; PRIOR FILING DATE: 1998-03-30

; PRIOR APPLICATION NUMBER: 60/080105

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080107

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080165

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080194

; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/080327

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/080328

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/080333

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/080334

; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/081070

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081049

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081071

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081195

; PRIOR FILING DATE: 1998-04-08

; PRIOR APPLICATION NUMBER: 60/081203

; PRIOR FILING DATE: 1998-04-09

; PRIOR APPLICATION NUMBER: 60/081229

; PRIOR FILING DATE: 1998-04-09

; PRIOR APPLICATION NUMBER: 60/081955

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/081817

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/081819

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/081952

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/081838

; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/082568

; PRIOR FILING DATE: 1998-04-21

; PRIOR APPLICATION NUMBER: 60/082569

; PRIOR FILING DATE: 1998-04-21

; PRIOR APPLICATION NUMBER: 60/082704

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082804

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082700

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082797

; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082796

; PRIOR FILING DATE: 1998-04-23

; PRIOR APPLICATION NUMBER: 60/083336

; PRIOR FILING DATE: 1998-04-27

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/083392

; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 539 GGGPGAR 545

RESULT 7
US-09-978-697-183
; Sequence 183, Application US/09978697
; Patent No. US20020169284A1

;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Baker Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleon
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Kijavlin, Ivar J.
;; APPLICANT: Kuo, Sophia S.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Shelton, David L.
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; FILE REFERENCE: P2630PLC27
;; CURRENT APPLICATION NUMBER: US/09/978,697
;; CURRENT FILING DATE: 2001-10-16
;; PRIOR APPLICATION NUMBER: 09/918585
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27

APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630P1C9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
|||||
Db 539 GGGPGAR 545

RESULT 8
US-09-978-192A-183
; Sequence 183, Application US/09978192A
; Patent No. US2002017753A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon

;
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414

;
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 85.4% Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
| | | | |
Db 539 GGGPGAR 545

RESULT 10
US-09-978-189-183
; Sequence 183, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;

;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 85.4%; Score 35; DB 9; Length 713;

Best Local Similarity 85.7%; Pred. No. 2.7e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGGPGKR 7

|||||

Db 539 GGGPGAR 545

RESULT 9

US-09-999-832A-183

;; Sequence 183, Application US/09999832A

;; Publication No. US20020192706A1

GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Baker Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnovers, Luc
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleon
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Henspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.

;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Kljavin, Ivar J.
;; APPLICANT: Kuo, Sophia S.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James;
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Shelton, David L.
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2630PIC63
;; CURRENT APPLICATION NUMBER: US/09/999,832A
;; CURRENT FILING DATE: 2001-10-24
;; PRIOR APPLICATION NUMBER: 09/918585
;; PRIOR FILING DATE: 2001-07-30
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/064249
;; PRIOR FILING DATE: 1997-11-03
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066364
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077641
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078004
;; PRIOR FILING DATE: 1998-03-13
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078936
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079689
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079920
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31

; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 539 GGGPGAR 545

RESULT 11
US-10-174-590-416
; Sequence 416, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18

; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 416
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-416

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 539 GGGPGAR 545

RESULT 12
US-10-176-758-416
; Sequence 416, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 416
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-416

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 539 GGGPGAR 545

RESULT 13
US-10-175-737-416
; Sequence 416, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 416
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-416

APPLICANT: Paoni, Nicholas F.
APPLICANT: ROY, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC7
CURRENT APPLICATION NUMBER: US/09/978,189
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
 |||||
Db 539 GGGPGAR 545

RESULT 14

US-10-173-706-416
; Sequence 416, Application US/10173706
; Publication No. US2003002229A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 416
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-706-416

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
 |||||
Db 539 GGGPGAR 545

RESULT 15

US-10-175-738-416
; Sequence 416, Application US/10175738
; Publication No. US2003002229A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 416
; LENGTH: 713
; TYPE: PRT

; ORGANISM: Homo Sapien
US-10-175-738-416

Query Match 85.4%; Score 35; DB 9; Length 713;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
 |||||
Db 539 GGGPGAR 545

Search completed: February 4, 2003, 09:33:42
Job time : 20 secs

THIS PAGE BLANK (USPTO)

A:Gene: BDNF
 C:Superfamily: nerve growth factor beta chain
 C:Keywords: glycoprotein
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-150/Domain: propeptide #status predicted <PRO>
 F:151-269/Product: brain-derived neurotrophic factor #status predicted <MAT>
 F:143/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:163-230,208-259,218-261/Disulfide bonds: #status predicted

Query Match 90.2%; Score 37; DB 2; Length 269;
 Best Local Similarity 85.7%; Pred. No. 54;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
 |||||:|
 Db 52 GGGPGQR 58

RESULT 3
 G84325
 hypothetical protein Vngl737h [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84325
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: G84325
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-323 <SPO>
 A:Cross-references: GB:AE004437; NID:g10581198; PIDN:AAG19971.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: Vngl737H

Query Match 90.2%; Score 37; DB 2; Length 323;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
 |||||:|
 Db 38 GGGPGER 44

RESULT 4
 E70018
 sugar permease homolog yurN - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: E70018
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emmerson, P.T.; Encian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpst, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: E70018
 A:Status: preliminary;
 A:Molecule type: DNA

A:Residues: 1-292 <KUN>
 A:Cross-references: GB:299120; GB:AL009126; NID:g2635613; PIDN:CAB15249.1; PID:g26357
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yurN
 C:Superfamily: inner membrane protein ugpA

Query Match 87.8%; Score 36; DB 2; Length 292;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6
 |||||:
 Db 237 GGGPGK 242

RESULT 5
 G82139
 endopeptidase Clp (EC 3.4.21.92) chain P VC1922 [similarity] - Vibrio cholerae (stral
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Aug-2001
 C:Accession: G82139
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: G82139
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-200 <HEI>
 A:Cross-references: GB:AE004268; GB:AE003852; NID:g9656456; PIDN:AAF95070.1; GSPDB:GN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1922
 A:Map position: 1
 C:Superfamily: endopeptidase Clp chain P
 C:Keywords: hydrolase; serine proteinase
 F:103/Active site: Ser #status predicted
 F:128/Active site: His #status predicted

Query Match 85.4%; Score 35; DB 2; Length 200;
 Best Local Similarity 85.7%; Pred. No. 88;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
 |||||:
 Db 111 GGAPGKR 117

RESULT 6
 B71314
 endopeptidase Clp (EC 3.4.21.92) chain P1 [similarity] - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 02-Apr-2001
 C:Accession: B71314
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: B71314
 A:Status: preliminary;
 A:Molecule type: DNA
 A:Residues: 1-203 <COL>
 A:Cross-references: GB:AE001227; GB:AE000520; NID:g3322797; PIDN:AAC65495.1; PID:g332
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0507
 C:Superfamily: endopeptidase Clp chain P
 C:Keywords: hydrolase; serine proteinase

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 4, 2003, 08:50:51 ; Search time 40 seconds
(without alignments)
16.824 Million cell updates/sec

Title: US-09-706-690-1
Perfect score: 41
Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 38 | 92.7 | 189 | 2 A48834 | basic fibroblast g |
| 2 | 37 | 90.2 | 269 | 2 I51708 | brain-derived neur |
| 3 | 37 | 90.2 | 323 | 2 G84325 | hypothetical prote |
| 4 | 36 | 87.8 | 292 | 2 E70018 | sugar permease hom |
| 5 | 35 | 85.4 | 200 | 2 G82139 | endopeptidase Clp |
| 6 | 35 | 85.4 | 203 | 2 B71314 | endopeptidase Clp |
| 7 | 35 | 85.4 | 296 | 2 T12469 | hypothetical prote |
| 8 | 35 | 85.4 | 387 | 2 JC2032 | regulatory protein |
| 9 | 35 | 85.4 | 654 | 2 AG3522 | iron-regulated out |
| 10 | 34 | 82.9 | 76 | 2 T09262 | glycine-rich cell |
| 11 | 34 | 82.9 | 402 | 2 A45056 | osteogenic protein |
| 12 | 34 | 82.9 | 550 | 2 T34771 | signal recognition |
| 13 | 34 | 82.9 | 572 | 2 T12496 | hypothetical prote |
| 14 | 34 | 82.9 | 582 | 2 H95950 | probable ABC trans |
| 15 | 34 | 82.9 | 1870 | 2 S37671 | MHC class III hist |
| 16 | 34 | 82.9 | 1872 | 2 S36152 | MHC class III hist |
| 17 | 34 | 82.9 | 2142 | 2 B35098 | MHC class III hist |
| 18 | 33 | 80.5 | 115 | 2 AG3279 | hypothetical prote |
| 19 | 33 | 80.5 | 128 | 2 T35073 | probable phosphori |
| 20 | 33 | 80.5 | 136 | 2 T31545 | hypothetical prote |
| 21 | 33 | 80.5 | 138 | 2 T03374 | high mobility grou |
| 22 | 33 | 80.5 | 148 | 2 T26309 | hypothetical prote |
| 23 | 33 | 80.5 | 248 | 2 E84500 | hypothetical prote |
| 24 | 33 | 80.5 | 281 | 2 AH3012 | outer surface prot |
| 25 | 33 | 80.5 | 284 | 2 G98271 | hypothetical prote |
| 26 | 33 | 80.5 | 286 | 2 H70812 | hypothetical prote |
| 27 | 33 | 80.5 | 305 | 2 AB1321 | probable phosphotr |
| 28 | 33 | 80.5 | 312 | 2 A11329 | hypothetical prote |
| 29 | 33 | 80.5 | 312 | 2 A11700 | hypothetical prote |

| | | | | | |
|----|----|------|------|----------|--------------------|
| 30 | 33 | 80.5 | 324 | 2 H72657 | hypothetical prote |
| 31 | 33 | 80.5 | 338 | 2 A69125 | hypothetical prote |
| 32 | 33 | 80.5 | 342 | 2 S18649 | homeotic protein H |
| 33 | 33 | 80.5 | 359 | 2 T51903 | related to HLUJ pr |
| 34 | 33 | 80.5 | 366 | 1 A40056 | inhibin alpha chai |
| 35 | 33 | 80.5 | 403 | 2 T39697 | DNAJ protein - fis |
| 36 | 33 | 80.5 | 412 | 2 A84455 | hypothetical prote |
| 37 | 33 | 80.5 | 524 | 2 A59081 | acetolactate synth |
| 38 | 33 | 80.5 | 532 | 2 T35119 | probable aminotran |
| 39 | 33 | 80.5 | 564 | 2 C84456 | hypothetical prote |
| 40 | 33 | 80.5 | 627 | 1 JC6534 | protein kinase 1 (|
| 41 | 33 | 80.5 | 672 | 2 B84782 | probable receptor- |
| 42 | 33 | 80.5 | 705 | 2 T01137 | hypothetical prote |
| 43 | 33 | 80.5 | 789 | 2 T52067 | hypothetical prote |
| 44 | 33 | 80.5 | 844 | 2 T32608 | hypothetical prote |
| 45 | 33 | 80.5 | 1135 | 2 I61186 | alpha-7 integrin - |

ALIGNMENTS

RESULT 1
A48834
basic fibroblast growth factor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: A48834; S23636
R:Borja, A.Z.; Meijers, C.; Zeller, R.
Dev. Biol. 157, 110-118, 1993
A:Title: Expression of alternatively spliced bFGF first coding exons and antisense mR
A:Reference number: A48834; MUID:93246053; PMID:7683281
A:Accession: A48834
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-189 <BOR>
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:131000, NCBIP:131001)
R:Mitrani, E.; Gruenbaum, Y.; Shohat, H.; Ziv, T.
Development 109, 387-393, 1990
A:Title: Fibroblast growth factor during mesoderm induction in the early chick embryo
A:Reference number: S23636; MUID:90382254; PMID:2401202
A:Accession: S23636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 95-128 <MIT>
A:Cross-references: EMBL:X56804; NID:g62855; PIDN:CAAA0139.1; PID:g62856
C:Superfamily: fibroblast growth factor

Query Match 92.7%; Score 38; DB 2; Length 189;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGGPGKR 7
|||||:
Db 18 GGGPGRR 24

RESULT 2
I51708
brain-derived neurotrophic factor precursor - southern platyfish
C:Species: Xiphophorus maculatus (southern platyfish)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: I51708; S26673
R:Gotz, R.; Raulf, F.; Scharlt, M.
J. Neurochem. 59, 432-442, 1992
A:Title: Brain-derived neurotrophic factor is more highly conserved in structure and
A:Reference number: I51708; MUID:92333301; PMID:1629719
A:Accession: I51708
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-269 <GOT>
A:Cross-references: EMBL:X59942; NID:g65275; PIDN:CAAA2567.1; PID:g65276
C:Genetics:

Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 20 GGGPGLR 26

RESULT 12
T34771
signal recognition particle protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000
C:Accession: T34771
R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A:Reference number: 221557
A:Accession: T34771
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-550 <MUR>
A:Cross-references: EMBL:AL023797; PIDN:CAA19378.1; GSPDB:GN000070; SCOEDB:SC2E1.03
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: ffh; SCOEDB:SC2E1.03
C:Superfamily: signal recognition particle 54K protein

Query Match 82.9%; Score 34; DB 2; Length 550;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 483 GGGPGRQ 489

RESULT 13
T12496
hypothetical protein DKFZp434I091.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C:Accession: T12496; T46492
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: 217525
A:Accession: T12496
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-572 <POU>
A:Cross-references: EMBL:AL080174
A:Experimental source: adult testis; clone DKFZp434I091
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223029
A:Accession: T46492
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 456-572 <AAA>
A:Cross-references: EMBL:AL137328
A:Experimental source: adult testis; clone DKFZp434A2372
C:Genetics:
A:Note: DKFZp434I091.1; DKFZp434A2372.1

Query Match 82.9%; Score 34; DB 2; Length 572;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 105 GGGPGR 111

RESULT 14

H95950
Probable ABC transporter ATP-binding protein, consisting of 2 fused ATP-binding domain
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: H95950
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: H95950
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <KUP>
A:Cross-references: GB:AL591985; PIDN:CAC49272.1; PID:gl5140758; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymb
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB21206
A:Genome: plasmid
C:Superfamily: Escherichia coli ABC transporter ybhF; ATP-binding cassette homology

Query Match 82.9%; Score 34; DB 2; Length 582;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 310 GGGPGR 316

RESULT 15
S37671
MHC class III histocompatibility antigen HLA-B-associated protein 2 [similarity] - hu
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Sep-2000
C:Accession: S37671
R:Bougueleret, L.
submitted to the EMBL Data Library, August 1992
A:Reference number: S37671
A:Accession: S37671
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1870 <BOU>
A:Cross-references: EMBL:Z15025; NID:g29374; PID:g29375
C:Genetics:
A:Map position: 6p21.3
A:Introns: 38/2; 97/2; 129/3; 154/1; 202/1; 252/3; 279/2; 327/1; 357/2; 429/3; 588/1;
C:Superfamily: collagen alpha 1(IV) chain

Query Match 82.9%; Score 34; DB 2; Length 1870;
Best Local Similarity 85.7%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||
Db 1412 GGGPGR 1418

Search completed: February 4, 2003, 09:32:38
Job time : 49 secs

F:103/Active site: Ser #status predicted
F:128/Active site: His #status predicted

Query Match 85.4%; Score 35; DB 2; Length 203;
Best Local Similarity 85.7%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
IIIIII

Db 111 GGAQPKR 117

RESULT 7

Tl2469
hypothetical protein DKFzp564C1940.1 - human (fragment)

C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C:Accession: Tl2469

R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999

A:Reference number: Z17525

A:Accession: Tl2469

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-296 <POU>

A:Cross-references: EMBL:AL080164

A:Experimental source: fetal brain; clone DKFzp564C1940

C:Genetics:

A:Note: DKFzp564C1940.1

Query Match 85.4%; Score 35; DB 2; Length 296;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7

IIIIII

Db 122 GGGPGAR 128

RESULT 8

JC2032

regulatory protein AcyB2 - Streptomyces sp.

C:Species: Streptomyces sp.

C:Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 04-Mar-2000

C:Accession: JC2032

R:Arisawa, A.; Kawamura, N.; Tsunekawa, H.; Okamura, K.; Tone, H.; Okamoto, R.

Biosci. Biotechnol. Biochem. 57, 2020-2025, 1993

A:Title: Cloning and nucleotide sequences of two genes involved in the 4'-O-acetylation of

A:Reference number: JC2031; MUID:94122440; PMID:7764361

A:Accession: JC2032

A:Molecule type: DNA

A:Residues: 1-387 <ART>

A:Note: the source is designated as Streptomyces thermotolerans

C:Comment: This protein encodes a novel regulatory protein that activates cyb1 expression.

C:Superfamily: Streptomyces regulatory protein AcyB2

Query Match 85.4%; Score 35; DB 2; Length 387;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7

IIIIII

Db 165 GGGPGDR 171

RESULT 9

AG3522

iron-regulated outer membrane protein frpB BMEII10105 [imported] - Brucella melitensis (s

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AG3522

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

.; Mazur, M.; Golsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A:Reference number: AD3252; PMID:11756688

A:Accession: AG3522

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-654 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL53346.1; PID:gl7984235; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEII10105

A:Map position: II

Query Match 85.4%; Score 35; DB 2; Length 654;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7

IIIIII

Db 80 GGGPSKR 86

RESULT 10

T09262

glycine-rich cell wall protein EMB31 - white spruce

C:Species: Picea glauca (white spruce)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999

C:Accession: T09262

R:Dong, J.Z.; Dunstan, D.I.

submitted to the EMBL Data Library, June 1996

A:Description: Gene expression during somatic embryogenesis.

A:Reference number: Z16588

A:Accession: T09262

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-76 <DON>

A:Cross-references: EMBL:L47748; NID:gl350525; PID:gl350526

C:Genetics:

A:Gene: EMB31

Query Match 82.9%; Score 34; DB 2; Length 76;

Best Local Similarity 85.7%; Pred. No. 55;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7

IIIIII

Db 67 GGGPGGR 73

RESULT 11

A45056

osteogenic protein 2 precursor - human

N:Alternate names: bone morphogenetic protein 8; OP-2

C:Species: Homo sapiens (man)

C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999

C:Accession: A45056

R:Ozskaynak, E.; Schneegelsberg, P.N.; Jin, D.F.; Clifford, G.M.; Warren, F.D.; Drier,

J. Biol. Chem. 267, 25220-25227, 1992

A:Title: Osteogenic protein-2. A new member of the transforming growth factor-beta su

A:Reference number: A45056; MUID:93094231; PMID:1460021

A:Accession: A45056

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-402 <OZK>

A:Cross-references: GB:M97016; NID:gl89389; PIDN:AAB01360.1; PID:gl89390

A:Experimental source: hippocampus

A:Note: sequence extracted from NCBI backbone (NCBIP:120189)

C:Genetics:

A:Gene: GDB:BMP8; OP-2

A:Cross-references: GDB:I36392

C:Superfamily: inhibin

Query Match 82.9%; Score 34; DB 2; Length 402;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 07:05:52 : Search time 21 seconds
(without alignments)
13.825 Million cell updates/sec

Title: US-09-706-690-1

Perfect score: 41

Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|---------------|-------|----------------|--------|--------------|--------------------|
| 1 | 38 | 92.7 | 1279 | 1 SCAP_HUMAN | Q12770 homo sapien |
| 2 | 37 | 90.2 | 269 | 1 BONF_XIPMA | Q02193 xiphophorus |
| 3 | 36 | 87.8 | 130 | 1 Y4JI_RHISN | P55509 rhizobium s |
| 4 | 36 | 87.8 | 292 | 1 YURN_BACSU | O32155 bacillus su |
| 5 | 35 | 85.4 | 200 | 1 CLPP_VIBCH | Q9K96 vibrio chol |
| 6 | 35 | 85.4 | 203 | 1 CLP1_TREPA | O83520 treponema p |
| 7 | 35 | 85.4 | 218 | 1 RRAS_MOUSE | P10833 mus musculu |
| 8 | 35 | 85.4 | 741 | 1 RED1_HUMAN | P78563 homo sapien |
| 9 | 34 | 82.9 | 402 | 1 BNP8_HUMAN | P34820 homo sapien |
| 10 | 34 | 82.9 | 1193 | 1 DNB1_HSVT2 | Q9wrl7 herpesvirus |
| 11 | 34 | 82.9 | 2142 | 1 BAT2_HUMAN | P48634 homo sapien |
| 12 | 33 | 80.5 | 128 | 1 HIS3_STRCO | Q9S2ul streptomyce |
| 13 | 33 | 80.5 | 312 | 1 MRW_LISIN | Q929x6 listeria in |
| 14 | 33 | 80.5 | 312 | 1 MRW_LISHO | O8Y517 listeria mo |
| 15 | 33 | 80.5 | 342 | 1 HXD9_HUMAN | P28356 homo sapien |
| 16 | 33 | 80.5 | 366 | 1 IHA_RAT | P17490 rattus norv |
| 17 | 33 | 80.5 | 426 | 1 SHAT_HUMAN | O15105 homo sapien |
| 18 | 33 | 80.5 | 494 | 1 NPPI_MOUSE | P56183 mus musculu |
| 19 | 33 | 80.5 | 626 | 1 TESK_HUMAN | Q15569 homo sapien |
| 20 | 33 | 80.5 | 627 | 1 TESK_MOUSE | O70146 mus musculu |
| 21 | 33 | 80.5 | 628 | 1 TESK_RAT | O63572 rattus norv |
| 22 | 33 | 80.5 | 731 | 1 SUPL_XENLA | O9pun2 xenopus lae |
| 23 | 33 | 80.5 | 1179 | 1 ITA7_MOUSE | Q61738 mus musculu |
| 24 | 33 | 80.5 | 1293 | 1 MLE_DROME | P24785 drosophila |
| 25 | 33 | 80.5 | 1739 | 1 CHD2_HUMAN | O14647 homo sapien |
| 26 | 33 | 80.5 | 2314 | 1 AKA6_RAT | Q9wvc7 rattus norv |
| 27 | 32 | 78.0 | 116 | 1 DTUX_DROME | O9v1k4 drosophila |
| 28 | 32 | 78.0 | 162 | 1 RL24_HORVU | P50888 ardeum vul |
| 29 | 32 | 78.0 | 163 | 1 RL24_ARATH | P38666 arabidopsis |
| 30 | 32 | 78.0 | 190 | 1 Y2H5_STRCO | P35925 streptomyce |
| 31 | 32 | 78.0 | 226 | 1 BTOD_SERMA | P36572 serratia ma |
| 32 | 32 | 78.0 | 275 | 1 HXDC_HUMAN | P35452 homo sapien |
| 33 | 32 | 78.0 | 299 | 1 NUCG_BOVIN | P38447 bos taurus |

RESULT 1

| ID | SCAP_HUMAN | STANDARD; | PRT; | 1279 AA. |
|----|---|-----------|------|----------|
| AC | Q12770; | | | |
| DT | 15-JUN-2002 (Rel. 41, Created) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last sequence update) | | | |
| DT | 15-JUN-2002 (Rel. 41, Last annotation update) | | | |
| DE | Sterol regulatory element binding protein cleavage-activating protein | | | |
| DE | (SREBP cleavage-activating protein) (SCAP). | | | |
| GN | SCAP OR KIAA0199. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | TISSUE=Hepatoma; | | | |
| RX | MEDLINE=20037503; PubMed=10570913; | | | |
| RA | Nakajima T., Hamakubo T., Kodama T., Inazawa J., Emi M.; | | | |
| RT | "Genomic structure and chromosomal mapping of the human sterol | | | |
| RT | regulatory element binding protein (SREBP) cleavage-activating | | | |
| RT | protein (SCAP) gene."; | | | |
| RL | J. Hum. Genet. 44:402-407(1999). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 3-1279 FROM N.A. | | | |
| RX | TISSUE=Bone marrow; | | | |
| RX | MEDLINE=96281124; PubMed=8724849; | | | |
| RA | Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.; | | | |
| RT | "Prediction of the coding sequences of unidentified human genes. V. | | | |
| RT | The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by | | | |
| RT | analysis of cDNA clones from human cell line KG-1."; | | | |
| RL | DNA Res. 3:17-24(1996). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 346-1279 FROM N.A. | | | |
| RX | TISSUE=Teratocarcinoma; | | | |
| RA | Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., | | | |
| RA | Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., | | | |
| RA | Wakatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M., | | | |
| RA | Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., | | | |
| RA | Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., | | | |
| RA | Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., | | | |
| RA | Ninomiya K., Iwayanagi T.; | | | |
| RT | "NEDO human cDNA sequencing project."; | | | |
| RL | Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases. | | | |
| RN | [4] | | | |
| RP | VARIANT ILE-798. | | | |
| RX | MEDLINE=20037509; PubMed=10570919; | | | |
| RA | Iwaki K., Nakajima T., Ota N., Emi M.; | | | |
| RT | "A common Ile796Val polymorphism of the human SREBP cleavage- | | | |
| RT | activating protein (SCAP) gene."; | | | |
| RL | J. Hum. Genet. 44:421-422(1999). | | | |
| CC | - - FUNCTION: Sterol sensor. Necessary for the proteolytic activation | | | |
| CC | of SREBPs by site-1 protease in the Golgi. | | | |
| CC | - - SUBUNIT: In a tight complex with SREBP (By similarity). | | | |
| CC | - - SUBCELLULAR LOCATION: Integral membrane protein. Moves from the | | | |
| CC | endoplasmic reticulum to the Golgi in the absence of sterols (By | | | |

p25377 saccharomyc
Q9v295 homo sapien
P32233 mus musculu
P43690 xenopus lae
Q38920 arabidopsis
P22007 saccharomyc
P49355 bos taurus
P49156 homo sapien
Q02293 rattus norv
O15232 homo sapien
P30333 bradyrhizob
P38456 marchantia

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 CC -1- FUNCTION: Cleaves peptides in various proteins in a process that
 CC requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a
 CC major role in the degradation of misfolded proteins (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins to small peptides in
 CC the presence of ATP and magnesium. Alpha-casein is the usual test
 CC substrate. In the absence of ATP, only oligopeptides shorter than
 CC five residues are cleaved (such as succinyl-Leu-Tyr-I-NHMEC; and
 CC Leu-Tyr-Leu-I-Tyr-Trp, in which the cleavage of the -Tyr-I-Leu-
 CC and -Tyr-I-Trp- bond also occurs).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE004268; AAF95070.1; -
 DR HSSP; P19245; ITRYF.
 DR MEROPS; S14.001; -
 DR TIGR; VC1922; -
 DR InterPro: IPR001907; CLP_protease.
 DR Pfam; PF00574; CLP_protease; 1.
 DR PRINTS; PR00127; CLPPROTEASEP.
 DR TIGRFAMS; TIGR00493; clpp; 1.
 DR PROSITE; PS00382; CLP_PROTEASE_HIS; 1.
 DR PROSITE; PS00381; CLP_PROTEASE_SER; 1.
 KW Hydrolase; Serine protease; Complete proteome.
 FT ACT_SITE 103 103
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 200 AA; 22130 MW; 961403B346AF3681 CRC64;
 Query Match 85.4%; Score 35; DB 1; Length 200;
 Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGPGKR 7
 || ||||
 Db 111 GGAPGKR 117
 RESULT 6
 CLP1_TREPA STANDARD; PRT; 203 AA.
 ID CLP1_TREPA STANDARD; PRT; 203 AA.
 AC O83520;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ATP-dependent Clp protease proteolytic subunit 1 (EC 3.4.21.92)
 DE (Endopeptidase Clp 1).
 GN CLPP OR CLPP-1 OR TP0507.
 OS *Treponema pallidum*.
 OC Bacteria; Spirochaetales; Spirochaetaceae; *Treponema*.
 OX NCBI_TaxID=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=96332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
 RA McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,

RA Venter J.C.;
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis
 RT spirochete.";
 RL Science 281:375-388(1998).
 CC -1- FUNCTION: Cleaves peptides in various proteins in a process that
 CC requires ATP hydrolysis. Has a chymotrypsin-like activity. Plays a
 CC major role in the degradation of misfolded proteins (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins to small peptides in
 CC the presence of ATP and magnesium. Alpha-casein is the usual test
 CC substrate. In the absence of ATP, only oligopeptides shorter than
 CC five residues are cleaved (such as succinyl-Leu-Tyr-I-NHMEC; and
 CC Leu-Tyr-Leu-I-Tyr-Trp, in which the cleavage of the -Tyr-I-Leu-
 CC and -Tyr-I-Trp- bond also occurs).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S14.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001227; AAC65495.1; -
 DR HSSP; P19245; ITRYF.
 DR MEROPS; S14.001; -
 DR TIGR; TP0507; -
 DR InterPro: IPR001907; CLP_protease.
 DR Pfam; PF00574; CLP_protease; 1.
 DR PRINTS; PR00127; CLPPROTEASEP.
 DR TIGRFAMS; TIGR00493; clpp; 1.
 DR PROSITE; PS00382; CLP_PROTEASE_HIS; 1.
 DR PROSITE; PS00381; CLP_PROTEASE_SER; 1.
 KW Hydrolase; Serine protease; Complete proteome.
 FT ACT_SITE 103 103
 FT ACT_SITE 128 128 BY SIMILARITY.
 SQ SEQUENCE 203 AA; 52692B1AAB45ADB3 CRC64;
 Query Match 85.4%; Score 35; DB 1; Length 203;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGPGKR 7
 || ||||
 Db 111 GGAPGKR 117
 RESULT 7
 RRAS_MOUSE STANDARD; PRT; 218 AA.
 ID RRAS_MOUSE STANDARD; PRT; 218 AA.
 AC P10833;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Ras-related protein R-Ras (P23).
 GN RRAS.
 OS *Mus musculus* (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87078390; PubMed=3098437;
 RA Lowe D.G., Capon D.J., Delwart E., Sakaguchi A.Y., Naylor S.L.,
 RA Goeddel D.V.;
 RT "Structure of the human and murine R-ras genes, novel genes closely
 RT related to ras proto-oncogenes.";
 RL Cell 48:137-146(1987).
 CC -1- SUBCELLULAR LOCATION: INNER SURFACE OF PLASMA MEMBRANE POSSIBLY
 CC WITH ATTACHMENT REQUIRING AYLATION OF THE C-TERMINAL CYSTEINE
 CC (BY SIMILARITY WITH RAS).

```
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 13.6 kDa protein Y4J1.
GN Y4J1.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234.
CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: NONE OBVIOUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000079; AAB91721.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 130 AA; 13606 MW; 90817531FDF83463 CRC64;

Query Match 87.8%; Score 36; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPKG 6
Db 44 GGGPKG 49

RESULT 4
YURN_BACSU
ID YURN_BACSU STANDARD; PRT; 292 AA.
AC O32155;
AT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical ABC transporter permease protein yurn.
GN YURN.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriest R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bouilliet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.-M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
```


Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||

Db 46 GGGPGRK 52

RESULT 9

BMP8_HUMAN STANDARD; PRT; 402 AA.

AC P34820; Q9NUF0;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone morphogenetic protein 8 precursor (BMP-8) (Osteogenic protein 2) (OP-2).
GN BMP8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=93094231; PubMed=1460021;
RA Oezkaynak E., Schneegelsberg P.N.J., Jin D.F., Clifford G.M.,
Warren F.D., Drier E.A., Oppermann H.;
RT "Osteogenic protein-2. A new member of the transforming growth
factor-beta superfamily expressed early in embryogenesis.";
RT J. Biol. Chem. 267:25220-25227(1992).
RN [2]
RP SEQUENCE FROM N.A.

RL Ellington A.;
RA Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE
CC OSTEOINDUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF
CC EPITHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION
CC AND BONE HOMEOSTASIS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M97016; AAB01360.1; -;
DR EMBL; AL033527; CAB75681.1; -;
DR PIR; A45056; A45056.
DR HSSP; P18075; 1BMP.
DR Genew; HGNC:1075; BMP8.
DR MIM; 602284; -;
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR Prodom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 263 POTENTIAL.
FT CHAIN 264 402 BONE MORPHOGENETIC PROTEIN 8.
FT DISULFID 301 367 BY SIMILARITY.
FT DISULFID 330 399 BY SIMILARITY.
FT DISULFID 334 401 BY SIMILARITY.
FT DISULFID 366 366 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 343 343 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 358 360 KPN -> MPD (IN REF. 2).
SQ SEQUENCE 402 AA; 44764 MW; AE2338D9F11B1DD9 CRC64;

Query Match 82.9%; Score 34; DB 1; Length 402;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||

Db 20 GGGPGLR 26
|||||

RESULT 10

DNBI_HSVT2 STANDARD; PRT; 1193 AA.

AC Q9WRL7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major DNA-binding protein.
GN DBP.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Daral G.;
RT "Structural organization of a conserved gene cluster of Tupaia
herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
processing and transport protein, and the major DNA binding protein.";
RL Virus Res. 60:123-136(1999).
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF084543; AAD42933.1; -;
DR InterPro; IPR000635; Viral_DNA_bind.
DR Pfam; PF00747; Viral_DNA_bp; 1.
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
FT ZN_FING 453 466 C2HC-TYPE.
SQ SEQUENCE 1193 AA; 129834 MW; 7BCA52E1B6FAC67D CRC64;

Query Match 82.9%; Score 34; DB 1; Length 1193;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
|||||

Db 563 GGGPGGR 569
|||||

RESULT 11

BAT2_HUMAN STANDARD; PRT; 2142 AA.

AC P48634;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Large proline-rich protein BAT2 (HLA-B-associated transcript 2).

```

CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE R-RAS SUB-FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21019; AAA40038.1; -;
DR HSPG; P01112; IPLL.
DR MGD; MGI:98179; Rras.
DR InterPro: IPR003577; GTPase_Ras.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001806; Ras_transfmrng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR SMART; PR00449; RASTRNSFRMNG.
DR TIGRfams; TIGR00231; small_gtp; 1.
KW GTP-binding; Prenylation; Lipoprotein.
FT NP_BIND 36 43 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 142 145 GTP (BY SIMILARITY).
FT DOMAIN 58 66 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 215 215 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 218 AA; 23764 MW; C1D32CE7904322E5 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 218;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 16 GGGGPR 22

RESULT 8
RED1_HUMAN
ID RED1_HUMAN STANDARD; PRT; 741 AA.
AC P78563; P78555; O00691; O00692; O00395; O00465;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Double-stranded RNA-specific editase 1 (EC 3.5.-.-) (dsRNA adenosine
DE deaminase) (RNA editing enzyme 1).
GN ADAR1 OR RED1 OR DRADA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID 9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Gerber A., O'Connell M.A., Keller W.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Mittaz L., Scott H.S., Rossier C., Antonarakis S.E.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE:97265373; PubMed:9111310;
RA Lai F., Chen C.X., Carter K.C., Nishikura K.;
RT "Editing of glutamate receptor B subunit ion channel RNAs by four
RT alternatively spliced DRADA2 double-stranded RNA adenosine
RT deaminases";
RL Mol. Cell. Biol. 17:2413-2424(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA Villard L., Tassone F., Haymowicz M., Welborn R., Gardiner K.;

```

```

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE:20289799; PubMed:10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- FUNCTION: EDITING OF THE MESSENGER RNAs FOR GLUTAMATE RECEPTOR
CC (GLUR) SUBUNITS BY SITE-SELECTIVE ADENOSINE DEAMINATION. EDITS
CC BOTH THE GLUR-B Q/R AND R/G SITES EFFICIENTLY BUT CONVERTS THE
CC ADENOSINE IN HOTSPOT1 MUCH LESS EFFICIENTLY.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; RED1-L/DRADA2B
CC (SHOWN HERE), RED1-S/DRADA2A AND DRADA2C; ARE PRODUCED BY
CC ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE ADENOSINE DEAMINASE EDITASE FAMILY.
CC -1- SIMILARITY: CONTAINS 2 DBM (DOUBLE-STRANDED RNA-BINDING) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U02120; AAB61686.1; -;
DR EMBL; U02121; AAB61687.1; -;
DR EMBL; X92227; CAA67611.1; -;
DR EMBL; X93833; CAA67762.1; -;
DR EMBL; U76420; AAC51240.1; -;
DR EMBL; U76421; AAC51241.1; -;
DR EMBL; U76422; AAC51242.1; -;
DR EMBL; AF001042; AAB58300.1; -;
DR EMBL; AL163301; CAB90493.1; -;
DR HSPG; Q91836; IDI2.
DR Genew; HGNC:226; ADAR1.
DR MIM; 601218; -;
DR InterPro: IPR002466; A_deamin.
DR InterPro: IPR001159; DS_RBD.
DR Pfam; PF00035; dsrm; 2.
DR Pfam; PF02137; A_deamin; 1.
DR SMART; SM00358; DSRM; 2.
DR PROSITE; P850141; A_DEAMIN_EDITASE; 1.
DR PROSITE; P850137; DS_RBD; 2.
KW mRNA processing; Hydrolase; zinc; RNA-binding; Repeat;
KW Alternative splicing.
FT DOMAIN 76 146 DBM 1.
FT DOMAIN 230 300 DBM 2.
FT METAL 394 394 ZINC (BY SIMILARITY).
FT ACT_SITE 396 396 BY SIMILARITY.
FT METAL 451 451 ZINC (BY SIMILARITY).
FT METAL 556 556 ZINC (BY SIMILARITY).
FT VARSPPLIC 466 505 MISSING (IN ISOFORM DRADA2A).
FT VARSPPLIC 713 741 ARLFTAFIKAGLAWKEKTEQDFSLTP -> VH (IN
FT ISOFORM DRADA2C).
FT CONFLICT 30 30 G -> A (IN REF. 4).
FT CONFLICT 423 423 R -> E (IN REF. 4).
FT CONFLICT 475 475 V -> L (IN REF. 4).
SQ SEQUENCE 741 AA; 80763 MW; 02B563414DD59C20 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 741;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;

```

```

Db      8 GGGPKR 13

RESULT 13
MRAW_LISIN
ID MRAW_LISIN STANDARD; PRT; 312 AA.
AC Q529X6; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR LIN2147.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL591982; CAD00119.1; -
CC ListinList; LMO02041; -
CC InterPro; IPR002903; Bac.Metrnfrse.
CC Pfam; PF01795; Methyltransf_5; 1.
CC ProDom; PD004685; Bac.Metrnfrse; 1.
CC TIGRFAMs; TIGR00006; UPF0117; 1.
CC Transferase; Methyltransferase; Complete proteome.
CC SEQUENCE 312 AA; 35429 MW; 44258BCC0F48BC32 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 312;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPKR 7
|||
Db 198 GGHPGKR 204

RESULT 14
MRAW_LISMO
ID MRAW_LISMO STANDARD; PRT; 312 AA.
AC Q8Y5L7; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR LMO2041.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -!- SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL596171; CAC97377.1; -
CC ListinList; LIN02147; -
CC InterPro; IPR002903; Bac.Metrnfrse.
CC Pfam; PF01795; Methyltransf_5; 1.
CC ProDom; PD004685; Bac.Metrnfrse; 1.
CC TIGRFAMs; TIGR00006; UPF0117; 1.
CC Transferase; Methyltransferase; Complete proteome.
CC SEQUENCE 312 AA; 35583 MW; F04715A45843A233 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 312;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGPKR 7
|||
Db 198 GGHPGKR 204

RESULT 15
HXD9_HUMAN
ID HXD9_HUMAN STANDARD; PRT; 342 AA.
AC P28356;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein Hox-D9 (Hox-4C) (Hox-5.2).
GN HOXD9 OR HOX4C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SOURCE FROM N.A.
RC TISSUE=Spinal cord;
RX MEDLINE=92097538; PubMed=1756725;
RA Zappavigna V., Renucci A., Izpisua-Belmonte J.-C., Urier G.,
RA Peschle C., Duboule D.;
RT "HOX4 genes encode transcription factors with potential auto- and
RT cross-regulatory capacities.";
RL EMBO J. 10:4177-4187(1991).
RN [2]
RN SEQUENCE OF 264-342 FROM N.A.
RX MEDLINE=89306602; PubMed=2568311;

```

GN BAT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=90192810; PubMed=2156268;
RA Banerji J., Sands J., Strominger J.L., Spies T.;
RT "A gene pair from the human major histocompatibility complex encodes
RT large proline-rich proteins with multiple repeated motifs and a
RT single ubiquitin-like domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2374-2378(1990).
RN [2]
RP SEQUENCE OF 1-1860 FROM N.A.
RX MEDLINE=93272029; PubMed=8499947;
RA Iris F.J.M., Bougueleret L., Prieur S., Caterina D., Primas G.,
RA Perrot V., Jurka J., Rodriguez-Tome P., Claverie J.-M., Dausset J.,
RA Cohen D.;
RT "Dense Alu clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment.";
RL Nat. Genet. 3:137-145(1993).
CC -1- FUNCTION: UNKNOWN.
CC -1- TISSUE SPECIFICITY: LIMITED TO CELL-LINES OF LEUKEMIC ORIGIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33509; AAA35585.1; -;
DR EMBL; M33518; AAA35586.1; -;
DR EMBL; M33512; AAA35586.1; JOINED.
DR EMBL; 215025; CAA78744.1; -;
DR PIR; B35098; B35098.
DR PIR; S36152; S36152.
DR Genev; HGNC:13918; BAT2.
DR MIM; 142580; -;
KW Repeat.
FT DOMAIN 519 POLY-PRO.
FT 524 GLN-RICH.
FT 636 POLY-PRO.
FT 657 POLY-PRO.
FT 684 POLY-PRO.
FT 699 POLY-PRO.
FT 704 POLY-PRO.
FT 814 POLY-PRO.
FT 821 POLY-GLY.
FT 1340 POLY-GLY.
FT 1398 POLY-GLY.
FT 1403 POLY-PRO.
FT 1436 POLY-PRO.
FT 1442 POLY-PRO.
FT 1982 POLY-PRO.
FT 1991 POLY-PRO.
FT 1795 4 X 57 AA TYPE A REPEATS.
FT 41 1-1.
FT REPEAT 41 95 1-1.
FT REPEAT 98 154 1-2.
FT REPEAT 281 337 1-3.
FT REPEAT 1740 1795 1-4.
FT 549 2 X TYPE B REPEATS.
FT 337 549 2-1.
FT REPEAT 337 418 2-1.
FT REPEAT 476 549 2-2.
FT 1899 3 X 50 AA TYPE C REPEATS.
FT DOMAIN 1899 2089 3 X 50 AA TYPE C REPEATS.
FT REPEAT 1899 1948 3-1.
FT REPEAT 1965 2014 3-2.
FT REPEAT 2040 2089 3-3.
FT REPEAT 57 57 R -> A (IN REF. 2).
FT REPEAT 109 109 R -> S (IN REF. 2).
FT CONFLICT 414 414 P -> PPHRGPAAGWGPP (IN REF. 2).
FT CONFLICT 532 532 T -> K (IN REF. 2).
FT CONFLICT 682 682 Q -> K (IN REF. 2).
FT CONFLICT 730 730 E -> D (IN REF. 2).
FT CONFLICT 750 750 L -> R (IN REF. 2).
FT CONFLICT 834 834 A -> T (IN REF. 2).
FT CONFLICT 1035 1035 G -> A (IN REF. 2).

FT CONFLICT 1068 1068 M -> L (IN REF. 2).
FT CONFLICT 1285 1285 P -> R (IN REF. 2).
FT CONFLICT 1400 1400 G -> A (IN REF. 2).
FT CONFLICT 1611 1611 T -> S (IN REF. 2).
FT CONFLICT 1729 1729 G -> A (IN REF. 2).
SQ SEQUENCE 2142 AA; 227840 MW; 32DDF16B9B52420A CRC64;

Query Match 82.9%; Score 34; DB 1; Length 2142;
Best Local Similarity 85.7%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 1401 GGGPGGR 1407
|||||

RESULT 12
HIS3_STRCO
ID HIS3_STRCO STANDARD; PRT; 128 AA.
AC O9S2UL;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) (PRA-CH).
GN HIS1 OR SCO2044 OR SC4G6.13C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-AMP + H(2)O -> 1-(5-
CC phosphoribosyl)-5-[(5-
CC phosphoribosylamino)methylideneamino]imidazole-4-carboxamide.
CC -1- PATHWAY: Histidine biosynthesis; third step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PRA-CH FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL096884; CAB51436.1; -;
DR InterPro; IPR002496; PRA-CH.
DR Pfam; PF01502; PRA-CH; 1.
DR -ProDom; PD002610; PRA-CH; 1.
KW Histidine biosynthesis; Hydrolase; Complete proteome.
SQ SEQUENCE 128 AA; 13834 MW; 3161011CC2990A82 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 128;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGK 6
|||||

RA Oliver G., Sidell N., Fiske N., Heinzmann C., Mohandas T.,
RT Sparkes R.S., de Robertis E.M.;
RA "Complementary homeo protein gradients in developing limb buds.";
RL Genes Dev. 3:641-650(1989).
RN [3]
RP SEQUENCE OF 275-340 FROM N.A.
RX MEDLINE=90098876; PubMed=2574852;
RA Acampora D., D'Esposito M., Faiella A., Pannese M., Migliaccio E.,
RA Morelli F., Stornaiuolo A., Nigro V., Simeone A., Boncinelli E.;
RT "The human HOX gene family.";
RL Nucleic Acids Res. 17:10385-10402(1989).
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING LIMB BUDS.
CC -!- SIMILARITY: BELONGS TO THE ASD-B HOMEBOX FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59372; CAA42016.1; -;
DR EMBL; X15506; CAA33528.1; -;
DR PIR; S18649; S18649.
DR PIR; S05958; S05958.
DR PIR; A32830; A32830.
DR HSPP; P02834; I1881.
DR TRANSPAC; T01424; -;
DR Genew; HGNC:5140; HOXD9.
DR MIM; 142982; -;
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 115 149 GLY-RICH.
FT DOMAIN 121 130 POLY-GLY.
FT DOMAIN 165 178 SER/THR-RICH.
FT DNA_BIND 275 334 HOMEBOX.
FT CONFLICT 266 266 E -> A (IN REF. 2).
SQ SEQUENCE 342 AA; 35580 MW; 731981FE25C5ACD7 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 342;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGK 6
 |||||
Db 128 GGGPGR 133

Search completed: February 4, 2003, 09:30:08
Job time : 30 secs

THIS PAGE BLANK (USPTO)

```

Qy 1 GGGPGKR 7
Db 866 GGGPGKR 872

RESULT 2
Q9AUN1 PRELIMINARY; PRT; 254 AA.
AC Q9AUN1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 26.7 kDa protein.
GN OSJNBA0058E19.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., King L., Kirchoff K.A., de la Bastide M., Preston R.R.,
RA Nascimmento L.U., Vil M.D., Baker J.P., Miller B., Cunnius D.M.,
RA Kuit K.H., Rodriguez S., Santos L., Zutavern T., Baliya V.S.,
RA Shah R.S., Bahret A., Bal H.P., O'Shaughnessy A., Dedhia N.N.,
RA McCombie W.R.;
RT "Genomic Sequence For Oryza sativa, Nipponbare Strain, Chromosome X,
RT Clone OSJNBA0058E19, Complete Sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC083945; AAK13145.1; -.
KW Hypothetical protein.
SQ SEQUENCE 254 AA; 26729 MW; 3C41BCF5A075EBA4 CRC64;

Query Match 92.7%; Score 38; DB 10; Length 254;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 99 GGGPGRR 105

RESULT 3
Q93V64 PRELIMINARY; PRT; 314 AA.
AC Q93V64;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE B1080D07.24 protein (P0507H06.8 protein).
GN B1080D07.24 OR P0507H06.8.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:B1080D07.24";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0507H06.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003203; BAB64072.1; -.
DR EMBL; AP003144; BAB44129.1; -.
SQ SEQUENCE 314 AA; 34061 MW; F0A7EB73774C8564 CRC64;

```

```

Query Match 92.7%; Score 38; DB 10; Length 314;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 99 GGGPGRR 105

RESULT 4
Q8SOV8 PRELIMINARY; PRT; 382 AA.
AC Q8SOV8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE OJ1014_G12.16 protein.
GN OJ1014_G12.16.
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
RT clone:OJ1014_G12.16";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003372; BAB89082.1; -.
SQ SEQUENCE 382 AA; 41171 MW; 4F4CF66BB649A0D6 CRC64;

Query Match 92.7%; Score 38; DB 10; Length 382;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGPGKR 7
Db 181 GGGPGRR 187

RESULT 5
Q9AYC4 PRELIMINARY; PRT; 609 AA.
AC Q9AYC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Putative retroelement pol polyprotein.
GN OSJNBA0094H10.11.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NIPPONBARE;
RA Spiegel L.A., Nascimmento L.U., de la Bastide M., Kirchoff K.A.,
RA King L., Preston R.R., Vil M.D., Baker J.P., Miller B., Zutavern T.,
RA Rodriguez S., Santos L., Kuit K.H., Cunnius D.M., Baliya V.S.,
RA Shah R.S., Bahret A., Bal H.P., O'Shaughnessy A., Dedhia N.N.,
RA McCombie W.R.;
RT "Genomic Sequence For Oryza sativa, Nipponbare Strain, Chromosome X,
RT Clone OSJNBA0058E19, complete sequence.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC080019; AAK13114.1; -.
KW Polyprotein.
SQ SEQUENCE 609 AA; 65261 MW; F940401C56B605B7 CRC64;

Query Match 92.7%; Score 38; DB 10; Length 609;

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 08:08:45 ; Search time 81 seconds
(without alignments)
17.807 Million cell updates/sec

Title: US-09-706-690-1
Perfect score: 41
Sequence: 1 GGGPGKR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|-----------|--------------------|
| 1 | 41 | 100.0 | 927 | 16 Q981V0 | Q981v0 rhizobium l |
| 2 | 38 | 92.7 | 254 | 10 Q9AUN1 | Q9aun1 oryza sativ |
| 3 | 38 | 92.7 | 314 | 10 Q93V64 | Q93v64 oryza sativ |
| 4 | 38 | 92.7 | 382 | 10 Q8S0V8 | Q8s0v8 oryza sativ |
| 5 | 38 | 92.7 | 609 | 10 Q9AVC4 | Q9ayc4 oryza sativ |
| 6 | 38 | 92.7 | 729 | 12 Q9DWC1 | Q9dwc1 rat cytoleg |
| 7 | 38 | 92.7 | 905 | 4 Q8WUAI | Q8wuai homo sapien |
| 8 | 37 | 90.2 | 323 | 17 Q9HP99 | Q9hp99 halobacteri |
| 9 | 37 | 90.2 | 384 | 5 Q8SYB8 | Q8syb8 drosophila |
| 10 | 37 | 90.2 | 875 | 6 Q46G06 | Q46g06 bos taurus |
| 11 | 36 | 87.8 | 59 | 2 Q9L9U5 | Q9l9u5 enterococcu |
| 12 | 36 | 87.8 | 121 | 4 Q8WVB2 | Q8wvb2 homo sapien |
| 13 | 36 | 87.8 | 172 | 4 Q96CG4 | Q96cg4 homo sapien |
| 14 | 36 | 87.8 | 533 | 10 Q9M7B9 | Q9m7b9 triglochin |
| 15 | 36 | 87.8 | 540 | 10 Q9M7C0 | Q9m7c0 triglochin |
| 16 | 36 | 87.8 | 802 | 10 Q8S2A7 | Q8s2a7 oryza sativ |

| | | | | | |
|----|----|------|------|-----------|---------------------|
| 17 | 35 | 85.4 | 220 | 4 Q9HC42 | Q9hc42 homo sapien |
| 18 | 35 | 85.4 | 231 | 4 Q958B2 | Q958b2 homo sapien |
| 19 | 35 | 85.4 | 294 | 16 Q93JD8 | Q93jd8 streptomyce |
| 20 | 35 | 85.4 | 296 | 4 Q9UG33 | Q9ug33 homo sapien |
| 21 | 35 | 85.4 | 348 | 11 Q99M85 | Q99m85 mus musculu |
| 22 | 35 | 85.4 | 387 | 2 Q56076 | Q56076 streptomyce |
| 23 | 35 | 85.4 | 449 | 11 Q9JJE7 | Q9jje7 mus musculu |
| 24 | 35 | 85.4 | 470 | 5 Q9BLR3 | Q9blr3 leishmania |
| 25 | 35 | 85.4 | 512 | 5 Q95ZG9 | Q95zg9 chironomus |
| 26 | 35 | 85.4 | 517 | 5 Q95ZH0 | Q95zh0 chironomus |
| 27 | 35 | 85.4 | 535 | 5 Q9UIN0 | Q9uin0 chironomus |
| 28 | 35 | 85.4 | 554 | 5 Q76233 | Q76233 trypanosoma |
| 29 | 35 | 85.4 | 654 | 16 Q8YDS0 | Q8yds0 bruceella me |
| 30 | 35 | 85.4 | 699 | 4 Q43263 | Q43263 homo sapien |
| 31 | 35 | 85.4 | 708 | 2 Q93N90 | Q93n90 streptomyce |
| 32 | 34 | 82.9 | 76 | 10 Q40855 | Q40855 picea glauc |
| 33 | 34 | 82.9 | 91 | 5 Q9VF73 | Q9vfv3 drosophila |
| 34 | 34 | 82.9 | 117 | 2 Q53075 | Q53075 rhodobacter |
| 35 | 34 | 82.9 | 264 | 13 Q9PT21 | Q9pt21 oncorhynchu |
| 36 | 34 | 82.9 | 550 | 2 Q85676 | Q85676 streptomyce |
| 37 | 34 | 82.9 | 550 | 16 Q69874 | Q69874 streptomyce |
| 38 | 34 | 82.9 | 572 | 4 Q9UG31 | Q9ug31 homo sapien |
| 39 | 34 | 82.9 | 576 | 10 Q94Z29 | Q94z29 oryza sativ |
| 40 | 34 | 82.9 | 582 | 16 Q92V45 | Q92v45 rhizobium m |
| 41 | 34 | 82.9 | 645 | 4 Q95928 | Q95928 homo sapien |
| 42 | 34 | 82.9 | 773 | 5 Q95TR3 | Q95tr3 drosophila |
| 43 | 34 | 82.9 | 1050 | 5 Q9BN18 | Q9bn18 drosophila |
| 44 | 34 | 82.9 | 1190 | 5 Q9W2K4 | Q9w2k4 drosophila |
| 45 | 34 | 82.9 | 1254 | 4 Q96AY4 | Q96ay4 homo sapien |

ALIGNMENTS

RESULT 1

| | | | | | |
|--------|---|---|--------------|------|---------|
| Q981V0 | ID | Q981V0 | PRELIMINARY; | PRT; | 927 AA. |
| AC | Q981V0; | | | | |
| DT | 01-OCT-2001 | (TREMBLrel. 18, Created) | | | |
| DT | 01-OCT-2001 | (TREMBLrel. 18, Last sequence update) | | | |
| DT | 01-JUN-2002 | (TREMBLrel. 21, Last annotation update) | | | |
| DE | Argininosuccinate lyase. | | | | |
| GN | ML9226. | | | | |
| OS | Rhizobium loti (Mesorhizobium loti). | | | | |
| OG | Plasmid pMLa. | | | | |
| OC | Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; | | | | |
| OC | Phyllobacteriaceae; Mesorhizobium. | | | | |
| OX | NCBI_TaxID:381; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN-MAFF303099; | | | | |
| RX | MEDLINE:21082930; PubMed:11214968; | | | | |
| RA | Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., | | | | |
| RA | Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., | | | | |
| RA | Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., | | | | |
| RA | Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., | | | | |
| RA | Takeuchi C., Yamada M., Tabata S.; | | | | |
| RT | *Complete genome structure of the nitrogen-fixing symbiotic bacterium | | | | |
| RT | Mesorhizobium loti.*; | | | | |
| RL | DNA Res. 7:331-338(2000). | | | | |
| DR | EMBL: AP003015; BAB54609.1; -. | | | | |
| DR | InterPro: IPR000901; CPSase. | | | | |
| DR | InterPro: IPR000362; Fumarate_lyase. | | | | |
| DR | Pfam: PF00206; lyase_1; 1. | | | | |
| DR | PRINTS; PR00149; FUMRATELYASE. | | | | |
| DR | PROSITE; PS00867; CPSASE.2; UNKNOWN_1. | | | | |
| KW | Lyase; Plasmid; Complete proteome. | | | | |
| SQ | SEQUENCE 927 AA; 99125 MW; 64E4255459968AB5 CRC64; | | | | |

Query Match 100.0%; Score 41; DB 16; Length 927;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RE73487p.
GN CG7903.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Clapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Krocbleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY071654; AAL49276.1; -.
SQ SEQUENCE 384 AA; 40162 MW; B49CC6DA16F22278 CRC64;

Query Match 90.2%; Score 37; DB 5; Length 384;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
   |||||
Db 360 GGGPGQR 366

RESULT 10
O46606 PRELIMINARY; PRT; 875 AA.
AC O46606;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Phosphatidic acid-prefering phospholipase A1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98157935; PubMed=9488669;
RA Higgs H.N., Han M.H., Johnson G.E., Glomset J.A.;
RT "Cloning of a phosphatidic acid-prefering phospholipase A1 from
RT bovine testis.";
RL J. Biol. Chem. 273:5468-5477(1998).
DR EMBL; AF045022; AAC03019.1; -.
DR InterPro; IPR004177; DDHD_dom.
DR Pfam; PF02862; DDHD; 1.
SQ SEQUENCE 875 AA; 97576 MW; E1DR4FFD7DC75EB1 CRC64;

Query Match 90.2%; Score 37; DB 6; Length 875;
Best Local Similarity 85.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGKR 7
   |||||
Db 137 GGGPGR 143

RESULT 11
Q9L9U5 PRELIMINARY; PRT; 59 AA.
AC Q9L9U5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

```

```

DE Enterocin 1071B prepeptide.
GN ENT1071B.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Plasmid pEF1071.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BFE 1071; PLASMID=PEF1071;
RX MEDLINE=20208822; PubMed=10742203;
RA Balla E., Dicks L.M.T., Du Toit M., Van Der Merwe M.J.,
RA Holzapfel W.H.;
RT "Characterization and cloning of the genes encoding enterocin 1071A
RT and enterocin 1071B, two antimicrobial peptides produced by
RT Enterococcus faecalis BFE 1071.";
RL Appl. Environ. Microbiol. 66:1298-1304(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BFE 1071; PLASMID=PEF1071;
RA Balla E.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=FAIR-E 309;
RA Franz C.M.A.P., Grube A., Herrmann A., Abriouel H., Staerke J.,
RA Lombardi A., Tauscher B., Holzapfel W.H.;
RT "Biochemical and genetic characterization of the two-peptide
RT bacteriocin enterocin 1071 produced by Enterococcus faecalis FAIR-E
RT 309 and amendment of the peptide sequence previously reported for
RT enterocin 1071A.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164560; AAF63953.2; -.
DR EMBL; AY063485; AAL39165.1; -.
KW Plasmid.
SQ SEQUENCE 59 AA; 6482 MW; 6FEC7862F437C3C8 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6
   |||||
Db 23 GGGPGK 28

RESULT 12
Q8WVB2 PRELIMINARY; PRT; 121 AA.
AC Q8WVB2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 12.5 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018206; AAH18206.1; -.
KW Hypothetical protein.
SQ SEQUENCE 121 AA; 12495 MW; C58AE8BF1E1A11E CRC64;

Query Match 87.8%; Score 36; DB 4; Length 121;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6
   |||||

```

| | |
|--|--|
| OX | NCBI_TaxID:-9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE-COLON; |
| RA | Strausberg R.; |
| RL | Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.. |
| DR | EMBL; BC020987; AAH20987.1; - |
| DR | InterPro: IPR000731; HMGCR/patch_5TM. |
| DR | InterPro: IPR001950; TIF_SUI1. |
| DR | InterPro: IPR001680; WD40. |
| DR | Pfam: PF00400; WD40; 6. |
| DR | PRINTS; PR00320; GPROTEINBRPT. |
| DR | SMART; SM00320; WD40; 6. |
| DR | PROSITE; PS01566; SSD; 1. |
| DR | PROSITE; PS01118; SUI1.1; UNKNOWN_1. |
| DR | PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1. |
| DR | PROSITE; PS50082; WD_REPEATS_2; 1. |
| DR | PROSITE; PS50294; WD_REPEATS_REGION; 1. |
| DR | Hypothetical protein. |
| KW | |
| SQ | SEQUENCE 905 AA; 97916 MW; F03D15E98177B8E9 CRC64; |
| | |
| Query Match | 92.7%; Score 38; DB 4; Length 905; |
| Best Local Similarity | 85.7%; Pred. No. 2e+02; |
| Matches 6; Conservative 1; Mismatches 0; Indels 0; Caps 0; | |
| QY | 1 GGGPGKR 7 : |
| DB | 369 GGGPGR 375 |
| | |
| RESULT 8 | |
| Q9HP99 | PRELIMINARY; PRT; 323 AA. |
| ID | Q9HP99 |
| AC | Q9HP99 |
| DT | 01-WAR-2001 (TrEMBLrel. 16, Created) |
| DT | 01-WAR-2001 (TrEMBLrel. 16, Last sequence update) |
| DT | 01-WAR-2001 (TrEMBLrel. 16, Last annotation update) |
| DE | Vng1737h. |
| GN | VNG1737h. |
| OS | Halobacterium sp. (strain NRC-1). |
| OC | Archaea; Euryarchaeota; Halobacteria; Halobacteriales; |
| OC | Halobacteriaceae; Halobacterium. |
| OX | NCBI_TaxID=64091; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE=20504483; PubMed=11016950; |
| RA | Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., |
| RA | Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J., |
| RA | Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., |
| RA | Leitchauer B., Keller K., Cruz R., Danson M.J., Hough D.W., |
| RA | Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., |
| RA | Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H., |
| RA | Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., |
| RA | Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.; |
| RT | "Genome sequence of Halobacterium species NRC-1."; |
| RL | Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000). |
| DR | EMBL; AF005079; BAG19971.1; - |
| KW | Complete proteome. |
| SQ | SEQUENCE 323 AA; 33850 MW; BFDC77096BE6F49AA CRC64; |
| | |
| Query Match | 90.2%; Score 37; DB 17; Length 323; |
| Best Local Similarity | 85.7%; Pred. No. 1e+02; |
| Matches 6; Conservative 1; Mismatches 0; Indels 0; Caps 0; | |
| QY | 1 GGGPGKR 7 : |
| DB | 38 GGGPGR 44 |
| | |
| RESULT 9 | |
| Q8SYB8 | PRELIMINARY; PRT; 384 AA. |
| ID | Q8SYB8 |
| AC | Q8SYB8; |

Db 109 GGGPGK 114

RESULT 13

Q96CG4 PRELIMINARY; PRT; 172 AA.

AC Q96CG4; 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 18.2 kDa protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014255; AAH14255.1; -;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 172 AA; 18221 MW; E407F5F13ACD149A CRC64;

Query Match 87.8%; Score 36; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6

|||||

Db 160 GGGPGK 165

RESULT 14

Q9M7B9 PRELIMINARY; PRT; 533 AA.

AC Q9M7B9; 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE Cytochrome P450 CYP79E2 (Fragment).
GN Triglochin maritimum.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Juncaginaceae; Triglochin.
OX NCBI_TaxID=55501;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20223747; PubMed=10759528;
RA Nielsen J.S., Moller B.L.;
RT "Cloning and expression of cytochrome P450 enzymes catalyzing the
conversion of tyrosine to p-hydroxyphenylacetaldoxime in the
biosynthesis of cyanogenic glucosides in Triglochin maritima.";
RL Plant Physiol. 122:1311-1321(2000).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF140610; AAF66544.1; -;
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
KW Heme; Monooxygenase; Oxidoreductase.
FT NON_TER 1
SQ SEQUENCE 533 AA; 59828 MW; 3B89CA74B5EC8DC4 CRC64;

Query Match 87.8%; Score 36; DB 10; Length 533;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6

|||||

Db 222 GGGPGK 227

RESULT 15

Q9M7C0 PRELIMINARY; PRT; 540 AA.

AC Q9M7C0; 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Cytochrome P450 CYP79E1.
GN CYP79E1.
OS Triglochin maritimum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Juncaginaceae; Triglochin.
OX NCBI_TaxID=55501;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20223747; PubMed=10759528;
RA Nielsen J.S., Moller B.L.;
RT "Cloning and expression of cytochrome P450 enzymes catalyzing the
conversion of tyrosine to p-hydroxyphenylacetaldoxime in the
biosynthesis of cyanogenic glucosides in Triglochin maritima.";
RL Plant Physiol. 122:1311-1321(2000).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF140609; AAF66543.1; -;
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 540 AA; 60815 MW; ED490AFFCA5292D3 CRC64;

Query Match 87.8%; Score 36; DB 10; Length 540;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGPGK 6

|||||

Db 226 GGGPGK 231

Search completed: February 4, 2003, 09:31:45
Job time : 92 secs

THIS PAGE BLANK (USPTO)

```

COMMENT
GENENTECH INC
OS Artificial gene
OC Artificial sequence; Genes.
OS homo sapiens
PN JP 1982163352-A/2
PD 07-OCT-1982
PF 29-DEC-1981 JP 1981215948
PR 02-JAN-1981 US 81 222044
PI DEIBUTSUDO BUI GOODERU, KEIICHI ITAKURA, DENISU JII KURIIDO
PC C07C103/52;C07G7/00;C07H21/04;C12N1/00;C12N15/00, PC
C12P21/00/C12R1/125,
PC C12R1/19,C12R1/42,C12R1/425;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Location/Qualifiers
FT CDS 6..185
FT /product='human proinsulin analogue' FT
FT mat_peptide 9..179
FT /product='human proinsulin analogue'.
FEATURES
source
1..185
/organism='synthetic construct'
/db_xref='taxon:32630' 64 t
BASE COUNT 38 a 45 c 38 g 64 t
ORIGIN
Query Match 65.4%; Score 119.6; DB 6; Length 185;
Best Local Similarity 82.4%; Pred. No. 9.3e-27;
Matches 150; Conservative 0; Mismatches 29; Indels 3; Gaps 1;

Qy 1 ATGTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGTTGTC 60
Db 6 ATGTTGTCATCAGCACCTTTGGTGTCTCACCTCGTTGAAGCTTGTACCTGTTGTC 65

Qy 61 GGTGAAGTGTGTTTTTCTACACCCCGAAACCGTGGTCCGGGTAAACGTGGGATC 120
Db 66 GGTGAAGTGTGTTTTTCTACACCTCTTACACTCTGTCTGCTGGATC---CAAGCGTGGGATC 122

Qy 121 GTTCAACAATGCTGTACTAGTACTGTCTCTCTACAGCTGGAGAACTATTGTAACTAG 180
Db 123 GTTGAACAGTGTGCACCTTCTATCTGTCTCTTTACACAGCTTGAGAACTACTGTAACATA 182

Qy 181 TA 182
Db 183 TA 184

RESULT 2
I04050
LOCUS I04050 171 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 4 from Patent EP 0128042.
ACCESSION I04050
VERSION I04050.1 GI:591905
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 171)
AUTHORS Berman,C.H. and Ullrich,A.
TITLE A process for identifying or isolating a DNA sequence, and a DNA hybridization probe therefor
JOURNAL Patent: Ep 0128042-A2 4 12-DEC-1984;
FEATURES
source
Location/Qualifiers
1..171
/organism='unknown'
BASE COUNT 34 a 42 c 35 g 60 t
ORIGIN
Query Match 59.8%; Score 109.4; DB 6; Length 171;
Best Local Similarity 80.4%; Pred. No. 1.5e-23;
Matches 144; Conservative 0; Mismatches 26; Indels 9; Gaps 1;

Qy 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGTGTAGAGCTCTGTACCTGGTTGGGT 63
Db 1 TTCGTCATCAGCACCTTTGTGGTCTCACCTCGTTGAAGCTTTGTACCTTTGTTGGGT 60

Qy 64 GAAGTGTGTTTTTCTACACCCCGAAACCGTGGTCCGGGTAAACGTGGGATCGTT 123
Db 61 GAAGTGTGTTTTTCTACACTCTCACTCCCTAAGAC-----TCGTCGTAAGCGTGGGATCGTT 111

Qy 124 GAACAATGCTGTACTAGTACTGTCTCTCTACAGCTGGAGAACTATTGTAACTAGTA 182
Db 112 GAACAGTGTGTCACCTTCTATCTGTCTCTTTACCAAGCTTGAGAACTACTGTAACATA 170

RESULT 3
SYNHUMINA
LOCUS SYNHUMINA 171 bp DNA linear SYN 27-APR-1993
DEFINITION Synthetic human insulin B and mini-C chains using deactivated silica gel chromatography.
ACCESSION M31026 M25442
VERSION M31026.1 GI:208535
KEYWORDS insulin.
SOURCE Synthetic DNA.
ORGANISM Synthetic construct.
REFERENCE 1 (bases 1 to 171)
AUTHORS Sung,W.L., Hsiung,H.M., Brousseau,R., Michniewicz,J., Wu,R. and Narang,S.A.
TITLE Synthesis of the human insulin gene. Part II. Further improvements in the modified phosphotriester method and the synthesis of seventeen deoxyribonucleotide fragments constituting human insulin chains B and mini-CDNA
JOURNAL Nucleic Acids Res. 7 (8), 2199-2212 (1979)
MEDLINE 80101076
PUBMED 230464
FEATURES
source
Location/Qualifiers
1..171
/organism='synthetic construct'
/db_xref='taxon:32630' 46 t
BASE COUNT 34 a 47 c 44 g 46 t
ORIGIN
Query Match 59.2%; Score 108.4; DB 12; Length 171;
Best Local Similarity 80.5%; Pred. No. 3.1e-23;
Matches 140; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

Qy 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGGCGT 63
Db 1 TTTGTCATCAGCACCTTTGGTGTCTCACCTGGTGGAGGCTCTGTACCTGGTGTGGGG 60

Qy 64 GAAGTGTGTTTTTCTACACCCCGAAACCGTGGTGGTCCGGGTAAACGTGGGATCGTT 123
Db 61 GAAGTGTGTTTTTCTACACACCCCAAGACC---CGTCGTAAGCTTAAGCGTGGGATCTGTG 117

Qy 124 GAACAATGCTGTACTAGTACTGTCTCTCTACAGCTGGAGAACTATTGTAAAC 177
Db 118 GAGCAGTGTGCACCAAGCATCTGCTCCTCTACCAACTGGAGAACTACTGCAAC 171

RESULT 4
AF050524
LOCUS AF050524 165 bp DNA linear SYN 19-NOV-1998
DEFINITION Synthetic Homo sapiens proinsulin-like protein BKRA gene, complete cds.
ACCESSION AF050524
VERSION AF050524.1 GI:3127059
KEYWORDS
SOURCE Synthetic construct.
ORGANISM Synthetic construct.
REFERENCE 1 (bases 1 to 165)
AUTHORS Fang,D., Wang,Y., Zhou,Z., Zhai,C., Gu,Z. and Wang,Y.

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 06:59:20 ; Search time 2725 Seconds
(without alignments)
1954.427 Million cell updates/sec

Title: US-09-706-690-3
Perfect score: 183
Sequence: 1 atgttcgttaatcagcacct.....agaactattgtaactagtaa 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vl.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|--------------------|
| 1 | 119.6 | 65.4 | 185 | E00055 | E00055 DNA encodin |
| 2 | 109.4 | 59.8 | 171 | I04050 | I04050 Sequence 4 |
| 3 | 108.4 | 59.2 | 171 | 12 SYNHUMINA | M1026 Synthetic h |
| 4 | 107.6 | 58.8 | 165 | 12 AF050524 | AF050524 Synthetic |
| 5 | 105.8 | 57.8 | 168 | I04049 | I04049 Sequence 3 |
| 6 | 100.2 | 54.8 | 162 | AX452029 | AX452029 Sequence |
| 7 | 96.6 | 52.8 | 195 | A08468 | A08468 Synthetic g |
| 8 | 96.6 | 52.8 | 195 | A08469 | A08469 Synthetic g |
| 9 | 93 | 50.8 | 415 | 6 AR007413 | AR007413 Sequence |
| 10 | 93 | 50.8 | 523 | 6 AR007425 | AR007425 Sequence |
| 11 | 92.2 | 50.4 | 258 | 6 AR077833 | AR077833 Sequence |
| 12 | 89.8 | 49.1 | 523 | 6 AR007431 | AR007431 Sequence |
| 13 | 89 | 48.6 | 192 | 6 A08012 | A08012 Synthetic g |
| 14 | 89 | 48.6 | 192 | 6 A08013 | A08013 Synthetic g |
| 15 | 89 | 48.6 | 192 | 6 I12516 | I12516 Sequence 45 |
| 16 | 89 | 48.6 | 192 | 6 I12517 | I12517 Sequence 46 |
| 17 | 88.4 | 48.3 | 352 | 6 AR094063 | AR094063 Sequence |
| 18 | 88.2 | 48.2 | 415 | 6 AR007419 | AR007419 Sequence |
| 19 | 88.2 | 48.2 | 415 | 6 AR007423 | AR007423 Sequence |
| 20 | 88 | 48.1 | 354 | 6 AR094062 | AR094062 Sequence |
| 21 | 87.6 | 47.9 | 1064 | 6 A46291 | A46291 Sequence 1 |
| 22 | 86.8 | 47.4 | 523 | 6 AR007415 | AR007415 Sequence |
| 23 | 86.6 | 47.3 | 415 | 6 AR007417 | AR007417 Sequence |
| 24 | 86.6 | 47.3 | 415 | 6 AR007421 | AR007421 Sequence |
| 25 | 84.8 | 46.3 | 162 | 6 A59450 | A59450 Sequence 43 |
| 26 | 84.8 | 46.3 | 162 | 6 AX182201 | AX182201 Sequence |
| 27 | 84.8 | 46.3 | 162 | 6 AX452035 | AX452035 Sequence |
| 28 | 84 | 45.9 | 219 | 6 AR023853 | AR023853 Sequence |
| 29 | 84 | 45.9 | 219 | 6 I46879 | I46879 Sequence 70 |
| 30 | 82.6 | 45.1 | 182 | 6 A15939 | A15939 Synthetic g |
| 31 | 82.6 | 45.1 | 183 | 6 A15938 | A15938 Synthetic g |
| 32 | 82.6 | 45.1 | 297 | 6 AR023835 | AR023835 Sequence |
| 33 | 82.6 | 45.1 | 297 | 6 I46861 | I46861 Sequence 44 |
| 34 | 82.6 | 45.1 | 335 | 6 A23271 | A23271 plasmid pla |
| 35 | 82.6 | 45.1 | 372 | 6 A23282 | A23282 plasmid pla |
| 36 | 82.6 | 45.1 | 381 | 6 A05979 | A05979 Synthetic n |
| 37 | 82.6 | 45.1 | 384 | 6 A05985 | A05985 Synthetic n |
| 38 | 82.6 | 45.1 | 399 | 6 A05983 | A05983 Synthetic n |
| 39 | 82.6 | 45.1 | 405 | 6 A05989 | A05989 Synthetic n |
| 40 | 82.6 | 45.1 | 408 | 6 A23280 | A23280 plasmid pla |
| 41 | 82.6 | 45.1 | 409 | 6 AR007427 | AR007427 Sequence |
| 42 | 82.6 | 45.1 | 414 | 6 A05987 | A05987 Synthetic n |
| 43 | 82.6 | 45.1 | 418 | 6 A05044 | A05044 leader-insu |
| 44 | 82.6 | 45.1 | 420 | 6 A23274 | A23274 plasmid plu |
| 45 | 82.6 | 45.1 | 426 | 6 A05981 | A05981 Synthetic n |

ALIGNMENTS

RESULT 1
E00055
LOCUS E00055 185 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human proinsulin analogue.
ACCESSION E00055
VERSION E00055.1 GI:2168361
KEYWORDS JP 1982163352-A/2.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 185)
Deibutsudo,B.G., Keilichi,I. and Denisu,J.K.
HUMAN PROINSULIN AND ANALOGUE, MANUFACTURE BY MICROBIAL POLYPEPTIDE
TITLE DEVELOPMENT AND CONVERSION TO HUMAN INSULIN
JOURNAL Patent: JP 1982163352-A 2 07-OCT-1982;

```
/product="synthetic.insulin"
/protein_id="CAA00783.1"
/db_xref="GI:411643"
/translation="NSNGKFVNQHLGSHLVEALYLVCGERGFFYTPKTRGIVEQCC
TSICSLYQENYCN"

BASE COUNT      40 a      65 c      52 g      38 t
ORIGIN

Query Match      52.8%; Score 96.6; DB 6; Length 195;
Best Local Similarity 76.0%; Pred. No. 1.6e-19;
Matches 139; Conservative 0; Mismatches 29; Indels 15; Gaps 1;

QY 1 ATGTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGTC 60
Db 13 AAGTTCGTCAACAGCACCTGTGGGCTCGCACCTCGTGGAGGCCCTCTACCTGGTGTGC 72
QY 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACCTGGGCATC 120
Db 73 GGGAGCGCGGCTCTCTACACCCCAAGACC-----AAGCGGGGCATC 117
QY 121 GTTGAACAATGCTGTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAAC TAG 180
Db 118 GTGAGCAGTGCTGTAGTCCATCTGCTCCCTCTACACGCTCGAGAACTACTGCAACTAG 177
QY 181 TAA 183
Db 178 TAA 180

RESULT 8
A08469/c
LOCUS      A08469      195 bp      DNA      linear      PAT 06-SEP-1993
DEFINITION Synthetic gene for human A0-Arg-Des-B30 insulin, Reverse complement.
ACCESSION  A08469
VERSION    A08469.1 GI:411644
KEYWORDS   synthetic construct.
SOURCE     synthetic construct.
ORGANISM   artificial sequences.
REFERENCE  1 (bases 1 to 195)
AUTHORS    Vertesy L., Geisen, K., Riess, G.J. and Sauber, K.
TITLE      Insulin derivatives, process for their preparation, their use and pharmaceutical compositions containing them
JOURNAL    Patent: EP 0376156-A 6 04-JUL-1990;
           HOECHST AKTIENGESSELLSCHAFT
FEATURES   Location/Qualifiers
            source
            1..195
            /organism="synthetic construct"
            /db_xref="taxon:32630"

BASE COUNT      38 a      52 c      65 g      40 t
ORIGIN

Query Match      52.8%; Score 96.6; DB 6; Length 195;
Best Local Similarity 76.0%; Pred. No. 1.6e-19;
Matches 139; Conservative 0; Mismatches 29; Indels 15; Gaps 1;

QY 1 ATGTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGTC 60
Db 183 AAGTTCGTCAACAGCACCTGTGGGCTCGCACCTCGTGGAGGCCCTCTACCTGGTGTGC 124
QY 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACCTGGGCATC 120
Db 123 GGGAGCGCGGCTCTCTACACCCCAAGACC-----AAGCGGGGCATC 79
QY 121 GTTGAACAATGCTGTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAAC TAG 180
Db 78 GTGAGCAGTGCTGTAGTCCATCTGCTCCCTCTACACGCTCGAGAACTACTGCAACTAG 19
QY 181 TAA 183
Db 18 TAA 16
```

```
RESULT 9
A007413
LOCUS      A007413      415 bp      DNA      linear      PAT 04-DEC-1998
DEFINITION Sequence 14 from patent US 5750497.
ACCESSION  A007413
VERSION    A007413.1 GI:3966897
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 415)
AUTHORS    Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth. and Markussen,J.
TITLE      Acylated insulin
JOURNAL    Patent: US 5750497-A 14 12-MAY-1998;
FEATURES   Location/Qualifiers
            source
            1..415
            /organism="unknown"

BASE COUNT      117 a      89 c      86 g      123 t
ORIGIN

Query Match      50.8%; Score 93; DB 6; Length 415;
Best Local Similarity 75.1%; Pred. No. 2.2e-18;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;

QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGCGGT 63
Db 227 TTCGTTAACCACACACTGTGGGTTCTCACTTGGTTGAAGCTTTGTACTTGGTTGGT 286
QY 64 GAACGTGGTTTTTTTACACCCCGAAACCGGTGGTCCGGGTAAACCTGGCATCGTT 123
Db 287 GAAAGAGGTTTCTTCTACACTCCAAAGCTCTGACGACGCTAAGGT-----ATCGTT 337
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAAC TAG 180
Db 338 GAACAATGTTGTACTTCTATCTCTTTGTACCAATTTGGAACAACTACTGTAAC TAG 394

RESULT 10
A007425
LOCUS      A007425      523 bp      DNA      linear      PAT 04-DEC-1998
DEFINITION Sequence 32 from patent US 5750497.
ACCESSION  A007425
VERSION    A007425.1 GI:3966909
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 523)
AUTHORS    Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth. and Markussen,J.
TITLE      Acylated insulin
JOURNAL    Patent: US 5750497-A 32 12-MAY-1998;
FEATURES   Location/Qualifiers
            source
            1..523
            /organism="unknown"

BASE COUNT      158 a      104 c      101 g      160 t
ORIGIN

Query Match      50.8%; Score 93; DB 6; Length 523;
Best Local Similarity 75.1%; Pred. No. 2.2e-18;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;

QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGCGGT 63
Db 335 TTCGTTAACCACACACTGTGGGTTCTCACTTGGTTGAAGCTTTGTACTTGGTTGGT 394
QY 64 GAACGTGGTTTTTTTACACCCCGAAACCGGTGGTCCGGGTAAACCTGGCATCGTT 123
Db 395 GAAAGAGGTTTCTTCTACACTCCAAAGCTCTGACGACGCTAAGGT-----ATCGTT 445
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAAC TAG 180
```

TITLE Synthesis of Human Proinsulin Analog (BKRA) Gene and Its Expression
in Escherichia coli
JOURNAL Zhongguo Shengwu Xue Yue Yu Fenzi Shengwu Xuebao 14 (5), 518-524
(1998)
REFERENCE 2 (bases 1 to 165)
AUTHORS Fang,D., Wang,Y., Zhou,Z., Zhai,C., Gu,Z. and Wang,Y.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1998) Huadong Research Institute for Medical
Biotechnics, 293 East Zhongshan Road, Nanjing, Jiangsu 210002,
People's Republic of China
FEATURES
source Location/Qualifiers
1..165
/organism="synthetic construct"
/db_xref="taxon:32630"
1..165
/note="similar to Homo sapiens proinsulin"
/codon_start=1
/transl_table=11
/product="proinsulin-like protein BKRA"
/protein_id="AAC78289.1"
/db_xref="GI:3127060"
/translation="MFVNQLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSLC
SLIQLNRYCN"
BASE COUNT 35 a 45 c 37 g 48 t
ORIGIN
Query Match 58.8%; Score 107.6; DB 12; Length 165;
Best Local Similarity 80.9%; Pred. No. 5.6e-23;
Matches 144; Conservative 0; Mismatches 19; Indels 15; Gaps 1;
Qy 1 ATGTTCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGC 60
Db 1 ATGTTCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGC 60
Qy 61 GCTGAAGCTGTTTCTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGGATC 120
Db 61 GGTGAAGCTGGCTCTTTTACATCCGAAAC-----TAAGCGCGGTATC 105
Qy 121 GTTCAACAACTCTACTAGACTCTCTCTCTACACGCTGGAGAACTATTGTAAC 178
Db 106 GTTGAACAGTGTTCACCTCATCTGCTCCCTTTACCAGCTGGAGAACTACTGTAAC 163
RESULT 5
I04049
LOCUS I04049 168 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 3 from Patent EP 0128042.
ACCESSION I04049
VERSION I04049.1 GI:591904
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 168)
AUTHORS Berman,C.H. and Ullrich,A.
TITLE A process for identifying or isolating a DNA sequence, and a DNA
hybridization probe therefor
JOURNAL Patent: EP 0128042-A2 3 12-DEC-1984;
FEATURES
source Location/Qualifiers
1..168
/organism="unknown"
BASE COUNT 39 a 51 c 42 g 36 t
ORIGIN
Query Match 57.8%; Score 105.8; DB 6; Length 168;
Best Local Similarity 79.7%; Pred. No. 2.1e-22;
Matches 141; Conservative 0; Mismatches 27; Indels 9; Gaps 1;
Qy 4 TTCGTTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGGCGT 63
Db 1 TTTGTGAACCAACACCTGTGGGCTCACACCTGTGGAGAGCTCTCTACCTAGTGTGGCGG 60
Qy 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGTGGTCCGGGTAAACGTGGCATGTT 123

Db 61 GAACGAGGCTTCTTCTACACACCCAAAGACC-----CGCCGGAAGCGTGGCATTCGTG 111
Qy 124 GAACAATGCTGTACTAGCATCTGCTCTCTTACCAGCTGGAGAACTATTGTAAC TAG 180
Db 112 GAACAATGCTGTACCAGCATCTGCTCCCTCTTACCAGCTGGAGAACTACTGCAAC TAG 168
RESULT 6
AX452029
LOCUS AX452029 162 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 6 from Patent EPI211314.
ACCESSION AX452029
VERSION AX452029.1 GI:21712032
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Annibali,N.
TITLE Expression of a human insulin precursor in p. Pastoris
JOURNAL Patent: EP 1211314-A 6 05-JUN-2002;
LABORATORIOS BETA S.A. (AR)
FEATURES
source Location/Qualifiers
1..162
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PCR product corresponding to the complete human
insulin precursor"
BASE COUNT 39 a 48 c 39 g 36 t
ORIGIN
Query Match 54.8%; Score 100.2; DB 6; Length 162;
Best Local Similarity 78.5%; Pred. No. 1.2e-20;
Matches 139; Conservative 0; Mismatches 23; Indels 15; Gaps 1;
Qy 4 TTCGTTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGGCGT 63
Db 1 TTTGTGAACCAACACCTGTGGGCTCACACCTGTGGAGCTCTCTACCTAGTGTGGCGG 60
Qy 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGTGGTCCGGGTAAACGTGGCATGTT 123
Db 61 GAACGAGGCTTCTTCTACACACCCAAAGACC-----AAGCGTGGCATTCGTG 105
Qy 124 GAACAATGCTGTACTAGCATCTGCTCTCTTACCAGCTGGAGAACTATTGTAAC TAG 180
Db 106 GAACAATGCTGTACCAGCATCTGCTCCCTCTTACCAGCTGGAGAACTACTGCAAC TAG 162
RESULT 7
A08468
LOCUS A08468 195 bp DNA linear PAT 06-SEP-1993
DEFINITION Synthetic gene for human A0-Arg-Des-B30 insulin.
ACCESSION A08468
VERSION A08468.1 GI:411642
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 195)
AUTHORS Vertesy,L., Gelsen,K., Riess,G.J. and Sauber,K.
TITLE Insulin derivatives, process for their preparation, their use and
pharmaceutical compositions containing them
JOURNAL Patent: EP 0376156-A 5 04-JUL-1990;
HOECHST AKTIENGESELLSCHAFT
FEATURES
source Location/Qualifiers
1..195
/organism="synthetic construct"
/db_xref="taxon:32630"
1..177
/partial
/codon_start=1
/transl_table=11
CDS


```
BASE COUNT      38 a      50 c      64 g      40 t
ORIGIN

Query Match      48.6%; Score 89; DB 6; Length 192;
Best Local Similarity 73.8%; Pred. No. 4e-17;
Matches 135; Conservative 0; Mismatches 30; Indels 18; Gaps 1;

QY 1 ATGTTGCGTTAATCAGCACCTGTGCGGCTCTCACCTGTAGAAAGCTCTGTACCTGGTTTGC 60
   ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 180 AAGTTCGTCAACAGCACCTGTGCGGCTCGCACCTCTGTGGAGGCCCTCTACCTGGTGTGC 121
   ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GGTGAACGTGTTTTTTTCTACACCCGAAACCGGTGGTGGTCCGGGTAAACGTGGCATC 120
   || || || || || || || || || || || || || || || || || || || || ||
Db 120 GGGAGCGCGGCTTCTTCTACACCCCAAGACC-----AAGGGCATC 79
   || || || || || || || || || || || || || || || || || || || || ||
QY 121 GTTGAACAATGCTGTACTAGCATCTGTCTCTTACCAGCTGGAGAACTATTGTAAC TAG 180
   || || || || || || || || || || || || || || || || || || || || ||
Db 78 GTGGAGCAGTCTGTAGTCCATCTGTCTCCCTCTACCAGCTCGAGAACTACTGCAACTAG 19
   || || || || || || || || || || || || || || || || || || || || ||
QY 181 TAA 183
   |||
Db 18 TAA 16

RESULT 15
I12516
LOCUS      I12516              192 bp      DNA      linear      PAT 26-JUL-1995
DEFINITION Sequence 45 from patent US 5426036.
ACCESSION      I12516
VERSION      I12516.1 GI:909900
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 192)
AUTHORS      Koller,K.-P., Riess,G., Uhlmann,E. and Wallmeier,H.
TITLE      Processes for the preparation of foreign proteins in streptomycetes
JOURNAL      Patent: US 5426036-A 45 20-JUN-1995;
FEATURES
    source      Location/Qualifiers
    1..192
               /organism="unknown"

BASE COUNT      40 a      64 c      50 g      38 t
ORIGIN

Query Match      48.6%; Score 89; DB 6; Length 192;
Best Local Similarity 73.8%; Pred. No. 4e-17;
Matches 135; Conservative 0; Mismatches 30; Indels 18; Gaps 1;

QY 1 ATGTTGCGTTAATCAGCACCTGTGCGGCTCTCACCTGTAGAAAGCTCTGTACCTGGTTTGC 60
   ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 13 AAGTTCGTCAACAGCACCTGTGCGGCTCGCACCTCTGTGGAGGCCCTCTACCTGGTGTGC 72
   ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GGTGAACGTGTTTTTTTCTACACCCGAAACCGGTGGTGGTCCGGGTAAACGTGGCATC 120
   || || || || || || || || || || || || || || || || || || || || ||
Db 73 GGGAGCGCGGCTTCTTCTACACCCCAAGACC-----AAGGGCATC 114
   || || || || || || || || || || || || || || || || || || || || ||
QY 121 GTTGAACAATGCTGTACTAGCATCTGTCTCTTACCAGCTGGAGAACTATTGTAAC TAG 180
   || || || || || || || || || || || || || || || || || || || || ||
Db 115 GTGGAGCAGTCTGTAGTCCATCTGTCTCCCTCTACCAGCTCGAGAACTACTGCAACTAG 174
   || || || || || || || || || || || || || || || || || || || || ||
QY 181 TAA 183
   |||
Db 175 TAA 177
```

Search completed: February 4, 2003, 08:50:43
Job time : 2735 secs

PT type I diabetes, comprises the properties of greater insulin receptor
PT binding activity than proinsulin and less insulin receptor binding
XX activity than insulin -

Example 1; Page 7; 24pp; English.

XX The invention relates to a single-chain insulin analogue (SIA) compound
XX comprising the properties of greater insulin receptor binding activity
CC than proinsulin and less insulin receptor binding activity than insulin.
CC The SIA compound has the formula: B chain - X - A chain; B and A = human
CC insulin chains or functional analogues; and X = a joining peptide of 5 to
CC 18 amino acids. The SIA compound or a vector comprising a polynucleotide
CC encoding the SIA compound is used to treat type I diabetes mellitus. The
CC present sequence represents a SIA-1 DNA sequence.

XX Sequence 183 BP; 40 A; 44 C; 46 G; 53 T; 0 other;

Query Match 100.0%; Score 183; DB 24; Length 183;
Best Local Similarity 100.0%; Pred. No. 5.1e-51;
Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTC 60
|||||
Db 1 ATGTTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTC 60

Qy 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
|||||

Db 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
|||||

Qy 121 GTTGAACAATGCTCTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAACCTAG 180
|||||

Db 121 GTTGAACAATGCTCTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAACCTAG 180
|||||

Qy 181 TAA 183
|||

Db 181 TAA 183

RESULT 2

AAN20002
ID AAN20002 standard; DNA; 185 BP.

AC AAN20002;

XX 17-DEC-1992 (first entry)

DE Human proinsulin analog chimeric gene consisting of A, B and C chains.
KW Insulin; hormone; chimeric protein; chimeric gene; fusion protein;
ds.

XX Homo sapiens.

XX EP55945-A.

XX 14-JUL-1982.

XX 31-DEC-1981; 81EP-0306193.

XX 02-JAN-1981; 81US-0222044.

XX (GENE-) GENENTECH INC.

XX Goeddel DV, Kleid DG, Itakura K;

XX WPI: 1982-59776E/29 (59776E).

XX P-PSDB: AAP20002.GBS.

XX Human pro:insulin for conversion to insulin - prepd. by microbial
PT expression of chimeric gene

XX Disclosure; Fig 5; 47pp; English.

XX

CC This sequence encodes segments of a gene for expression of an
CC analog of human proinsulin differing from human proinsulin in the
CC amino acid sequence of the C bridging chain specified in AAP20002.
CC This DNA sequence has sticky ends of AATT at the start and
CC AGCT at the finish.

XX Sequence 185 BP; 38 A; 45 C; 38 G; 64 T; 0 other;

Query Match 63.6%; Score 116.4; DB 3; Length 185;
Best Local Similarity 81.3%; Pred. No. 6.9e-29;
Matches 148; Conservative 0; Mismatches 31; Indels 3; Gaps 1;

Qy 1 ATGTTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTC 60
|||||
Db 6 ATGTTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTC 65

Qy 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
|||||

Db 66 GGTGAACGTGGTTTTTCTACACCTCTTAAGACTCGTCTGGGATC---CAAGCTGGGCATC 122
|||||

Qy 121 GTTGAACAATGCTCTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAACCTAG 180
|||||

Db 123 GTTGAACAGTGTTCACCTCTCTATCTGCTCTTTACCAGCTTGAGAACTACTGTAACTAA 182
|||||

Qy 181 TA 182
||

Db 183 TA 184

RESULT 3

AAT45974
ID AAT45974 standard; cDNA; 201 BP.

XX AAT45974;

XX 16-MAR-1997 (first entry)

XX DNA encoding Met-Arg-Met-single chain insulin.

XX Single chain insulin; diabetes mellitus; gene therapy; ss.

XX Synthetic.

XX EP741188-A2.

XX 06-NOV-1996.

XX 03-MAY-1996; 96EP-0303133.

XX 05-MAY-1995; 95US-0435762.

XX 05-MAY-1995; 95US-0435503.

XX (ELIL) LILLY & CO ELI.

XX Chance RE, Dimarchi RD, Hoffmann JA, Long HB, Miller AR;
PI Hoffman JA;

XX WPI: 1996-487391/49.

XX Single chain insulin polypeptide(s) - used for treating diabetes

XX Disclosure; Page 18; 22pp; English.

XX A nucleotide sequence (AAT45974) codes for Met-Arg-Met-single chain
CC insulin (AAN06809). The Met-Arg-Met N-terminal sequence can be
CC removed by CNBr cleavage to produce a single chain polypeptide
CC (AAN06807) that comprises the insulin B-chain joined to the A-chain
CC via a bridging peptide. The nucleotide sequence can be used to
CC produce recombinant single chain insulin, which has high
CC bioactivity, in transformed host cells. Gene therapy methods (see
CC also AAT45975) allow for the in vivo prodn. of the polypeptide under
CC glucose-regulated control for the treatment of diabetes.

XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 06:42:58 ; Search time 268 seconds
(without alignments)
1537.745 Million cell updates/sec

Title: US-09-706-690-3
Perfect score: 183
Sequence: 1 atgttgcttaacgcacct.....agaactattgtaactagtaa 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

| Database : | | | | N_Geneseq_101002.* |
|------------|--|-----|-------|---|
| 1: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1980.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 2: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1981.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 3: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1982.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 4: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1983.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 5: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1984.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 6: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1985.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 7: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1986.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 8: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1987.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 9: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1988.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 10: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1989.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 11: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1990.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 12: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1991.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 13: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1992.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 14: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1993.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 15: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1994.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 16: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1995.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 17: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1996.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 18: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1997.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 19: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1998.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 20: | /SID22/gcgdata/geneseq/geneseqn-embl/NA1999.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 21: | /SID22/gcgdata/geneseq/geneseqn-embl/NA2000.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 22: | /SID22/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 23: | /SID22/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 24: | /SID22/gcgdata/geneseq/geneseqn-embl/NA2002.DAT.* | 183 | 100.0 | Single-chain insulin analogue (SIA)-1 DNA sequence. |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|------------|-------|-------------|--------------|---|
| Result No. | Score | Query Match | Length DB ID | Description |
| 1 | 183 | 100.0 | 183 24 | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| 2 | 116.4 | 63.6 | 185 3 | Human proinsulin a |
| 3 | 113.4 | 62.0 | 201 17 | DNA encoding Met-A |
| 4 | 110.4 | 60.3 | 264 17 | Gene for single ch |
| 5 | 105.8 | 57.8 | 189 17 | DNA encoding singl |
| 6 | 100.4 | 54.9 | 171 21 | DNA encoding RGD-p |
| 7 | 93 | 50.8 | 415 16 | Yeast signal/leade |
| 8 | 93 | 50.8 | 523 16 | Mating factor alph |
| 9 | 92.2 | 50.4 | 258 17 | Proinsulin coding |

| | | | | | | |
|----|------|------|------|----|----------|---------------------|
| 10 | 91.2 | 49.8 | 550 | 18 | AAT69680 | DNA construct in p |
| 11 | 89.8 | 49.1 | 523 | 16 | AAQ86427 | Mating factor alph |
| 12 | 89 | 48.6 | 192 | 11 | AAQ04335 | Synthetic proinsul |
| 13 | 88.4 | 48.3 | 351 | 17 | AAT34670 | SOD-proinsulin hyb |
| 14 | 88.2 | 48.2 | 415 | 16 | AAQ86412 | Yeast signal/leade |
| 15 | 88.2 | 48.2 | 415 | 16 | AAQ86416 | Yeast signal/leade |
| 16 | 88 | 48.1 | 354 | 17 | AAT34669 | SOD-proinsulin hyb |
| 17 | 87.6 | 47.9 | 1063 | 16 | AAQ99460 | Glycosylphosphatid |
| 18 | 86.8 | 47.4 | 523 | 16 | AAQ86405 | Mating factor alph |
| 19 | 86.6 | 47.3 | 415 | 16 | AAQ86407 | Yeast signal/leade |
| 20 | 86.6 | 47.3 | 415 | 16 | AAQ86414 | Yeast signal/leade |
| 21 | 85.6 | 46.8 | 139 | 20 | AAQ86065 | PCR primer used to |
| 22 | 85.6 | 46.8 | 492 | 20 | AAQ86072 | Nucleotide sequenc |
| 23 | 85.6 | 46.8 | 550 | 22 | AAD13473 | YAP3-TA57 leader-N |
| 24 | 85.2 | 46.6 | 444 | 20 | AAQ86078 | Nucleotide sequenc |
| 25 | 85.2 | 46.6 | 552 | 20 | AAQ86076 | Nucleotide sequenc |
| 26 | 85.2 | 46.6 | 555 | 20 | AAQ86079 | Nucleotide sequenc |
| 27 | 85.2 | 46.6 | 708 | 20 | AAQ86090 | Nucleotide sequenc |
| 28 | 85.2 | 46.6 | 1191 | 20 | AAQ86091 | Nucleotide sequenc |
| 29 | 85.2 | 46.6 | 1344 | 20 | AAQ86092 | Nucleotide sequenc |
| 30 | 84.8 | 46.3 | 162 | 18 | AAT67150 | Ins1 double-chain |
| 31 | 84.2 | 46.0 | 409 | 16 | AAQ86423 | Yeast signal/leade |
| 32 | 84.2 | 46.0 | 660 | 17 | AAT07349 | pKV142 coding sequ |
| 33 | 84 | 45.9 | 219 | 17 | AAT10544 | N-terminally exten |
| 34 | 84 | 45.9 | 393 | 18 | AAT69678 | DNA construct in p |
| 35 | 84 | 45.9 | 550 | 19 | AAV42540 | PAK855 DNA sequenc |
| 36 | 84 | 45.9 | 550 | 19 | AAV41166 | DNA construct pAK7 |
| 37 | 83.8 | 45.8 | 159 | 17 | AAT45977 | Oligonucleotide fo |
| 38 | 82.6 | 45.1 | 188 | 8 | AAT71229 | Sequence of synthe |
| 39 | 82.6 | 45.1 | 297 | 17 | AAT10530 | S. cerevisiae expr |
| 40 | 82.6 | 45.1 | 335 | 13 | AAQ26521 | Plasmid pLAC202 fr |
| 41 | 82.6 | 45.1 | 372 | 13 | AAQ26527 | Plasmid pLAC05 frag |
| 42 | 82.6 | 45.1 | 381 | 10 | AAQ90970 | DNA sequence of th |
| 43 | 82.6 | 45.1 | 384 | 10 | AAQ90974 | DNA sequence of th |
| 44 | 82.6 | 45.1 | 399 | 10 | AAQ90973 | DNA sequence of th |
| 45 | 82.6 | 45.1 | 405 | 10 | AAQ90976 | DNA sequence of th |

ALIGNMENTS

RESULT 1

| | |
|----------|--|
| ABL58359 | ABL58359 standard; DNA; 183 BP. |
| ID | ABL58359 standard; DNA; 183 BP. |
| XX | ABL58359; |
| AC | ABL58359; |
| XX | |
| DT | 15-JUL-2002 (first entry) |
| XX | |
| DE | Single-chain insulin analogue (SIA)-1 DNA sequence. |
| XX | |
| KW | Single-chain insulin analogue; SIA; insulin; proinsulin; antidiabetic; |
| KW | SIA-1; gene therapy; ds. |
| XX | |
| OS | Synthetic. |
| OS | Homo sapiens. |
| XX | |
| PN | EP1193272-A1. |
| XX | |
| PD | 03-APR-2002. |
| XX | |
| PF | 13-SEP-2001; 2001EP-0121651. |
| XX | |
| PR | 02-OCT-2000; 2000KR-0058003. |
| PR | 07-NOV-2000; 2000US-0706690. |
| XX | (LEE/H) LEE H C. |
| XX | |
| PI | Lee HC, Kim S, Kim K, Shin H, Yoon J; |
| XX | |
| DR | WPI; 2002-373742/41. |
| XX | |
| PT | A single chain insulin analog (SIA) compound used in the treatment of |


```

SQ Sequence 201 BP; 40 A; 59 C; 55 G; 47 T; 0 other;
  Query Match      62.0%; Score 113.4; DB 17; Length 201;
  Best Local Similarity 79.0%; Pred. No. 7e-28;
  Matches 154; Conservative 0; Mismatches 26; Indels 15; Gaps 1;

QY 1 ATGTTCTGTTAATCAGCACCTGTGCGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGC 60
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 7 ATGTTTGTAAACCAACACCTGTGCGGCTCCACCTGGTAGAAGCTCTGTACCTGGTTGC 66
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GGTGAACGTGGTGTCTTCTACACCCCGAAGAC-----CGTGGTGGTCCG 105
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 67 GGTGAACGTGGTGTCTTCTACACCCCGAAGACCGAGCGCGCTGGCGCTCGAGGGTTCCCTG 126
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 106 GGTAAACGTGGCATCGTTGAACAATGCTGTACTAGCATCTGCTCTCTACCAAGTGGAG 165
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 127 CAGAACGTGGCATGTGTGAACAATGCTGTACCAGCATCTGCTCCCTGTACCAGCTGGAG 186
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 166 AACTATTGTAACCTAG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 187 AACTACTGCACTAG 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
AAT45975
ID AAT45975 standard; cDNA; 264 BP.
XX AC AAT45975;
XX DT 16-MAR-1997 (first entry)
XX DE Gene for single chain insulin with native signal peptide.
XX KW Single chain insulin; diabetes mellitus; gene therapy;
XX KW protein secretion; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT sig_peptide 1..72
FT /tag a
FT /note "human preproinsulin signal peptide"
FT mat_peptide 73..261
FT /tag b
FT /product single chain insulin
XX PN EP741188-A2.
XX PD 06-NOV-1996.
XX PF 03-MAY-1996; 96EP-0303133.
XX PR 05-MAY-1995; 95US-0435762.
XX PR 05-MAY-1995; 95US-0435503.
XX PA (ELIL ) LILLY & CO ELI.
XX PI Chance RE, Dimarchi RD, Hoffmann JA, Long HB, Miller AR;
PI Hoffman JA;
XX DR WPI; 1996-487391/49.
XX DR P-PSDB; AAW06810.
XX PT Single chain insulin polypeptide(s) - used for treating diabetes
XX PS Disclosure; Page 19; 22pp; English.
XX CC A synthetic gene (AAT45975) codes for a polypeptide (AAW06810)
XX CC consisting of the native human preproinsulin signal peptide and
XX CC a novel single chain insulin (see also AAW06807) in which the the
XX CC insulin B-chain is attached to the A-chain via a bridging peptide
XX CC derived from the C-chain. Inclusion of the signal sequence allows
XX CC the encoded construct to be secreted from transformed host cells.
```

```

CC The synthetic gene can be incorporated into a vector and utilised
CC in the gene therapy of type I or type II diabetes. The single
CC chain insulin can be produced in vivo under glucose-regulated
CC control.
SQ Sequence 264 BP; 44 A; 87 C; 76 G; 57 T; 0 other;
  Query Match      60.3%; Score 110.4; DB 17; Length 264;
  Best Local Similarity 78.6%; Pred. No. 7.6e-27;
  Matches 151; Conservative 0; Mismatches 26; Indels 15; Gaps 1;

QY 4 TTCGTTAATCAGCACCTGTGCGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGCGGT 63
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 73 TTTGTTAACCACACCTGTGCGGCTCCACCTGGTAGAAGCTCTGTACCTGGTTGCGGT 132
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GAACGTGGTGTCTTCTACACCCCGAAGAC-----CGGTGGTGGTCCGCGT 108
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 133 GAACGTGGTGTCTTCTACACCCCGAAGACGCGCGCTGGCGCTCGAGGGTTCCCTGCAG 192
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 109 AAACGTGGCATCGTTGAACAATGCTGTACTAGCATCTGCTCTCTACCAAGTGGAGAAC 168
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 193 AAGCGTGGCATGTGTGAACAATGCTGTACCAGCATCTGCTCCCTGTACCAGCTGGAGAAC 252
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 169 TATTGTAACCTAG 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 253 TACTGCAACTAG 264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
AAT45973
ID AAT45973 standard; DNA; 189 BP.
XX AC AAT45973;
XX DT 16-MAR-1997 (first entry)
XX DE DNA encoding single chain insulin polypeptide.
XX KW Single chain insulin; diabetes mellitus; gene therapy; ss.
XX OS Synthetic.
XX PN EP741188-A2.
XX PD 06-NOV-1996.
XX PF 03-MAY-1996; 96EP-0303133.
XX PR 05-MAY-1995; 95US-0435762.
XX PR 05-MAY-1995; 95US-0435503.
XX PA (ELIL ) LILLY & CO ELI.
XX PI Chance RE, Dimarchi RD, Hoffmann JA, Long HB, Miller AR;
PI Hoffman JA;
XX DR WPI; 1996-487391/49.
XX DR P-PSDB; AAW06807.
XX PT Single chain insulin polypeptide(s) - used for treating diabetes
XX PS Claim 9; Page 22; 22pp; English.
XX CC A nucleotide sequence (AAT45973) codes for a single chain peptide
XX CC (AAW06807) having high insulin bioactivity. Nucleotides encoding
XX CC such peptides (see also AAT45974-75) can be incorporated into vectors
XX CC and used to produce the peptides in transformed host cells. Gene
XX CC therapy methods allow for the in vivo prodn. of the peptides under
XX CC glucose-regulated control for the treatment of diabetes.
SQ Sequence 189 BP; 37 A; 57 C; 52 G; 43 T; 0 other;
  Query Match      57.8%; Score 105.8; DB 17; Length 189;
```

QY 4 TTCGTTAATCAGCACCTGTGGCGGCTCTCACCTGTGAGAGCTCTGTACCTGGTTGGCGGT 63
Db 1 TTCGTTAATCAGCACCTGTGGCGGCTCTCACCTGTGAGAGCTCTGTACCTGGTTGGCGGT 60
QY 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGTG 100
Db 61 GAACGTGGTTTTTCTACACCCCGAAACCGGTGCGG 97

RESULT 10
AAT69680
ID AAT69680 standard; DNA; 550 BP.
XX
AC AAT69680;
XX
DT 10-SEP-1997 (first entry)
XX
DE DNA construct in pAK749 for MI5 insulin precursor production.
XX
KW Recombinant protein; heterologous protein; MI5 insulin;
KW yeast; Saccharomyces cerevisiae; ds.
XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT CDS 115..492
FT sig_peptide /*tag= a
FT 115..177
FT /*tag= b
FT /product= YAP3 signal peptide
FT misc_RNA 178..307
FT /*tag= c
FT /product= LA19 leader peptide
FT misc_RNA 308..324
FT /*tag= d
FT /product= EEAEPK extension
FT mat_peptide 325..489
FT /*tag= e
FT /product= MI5 insulin precursor

XX
PN WO9722706-A1.
XX
PD 26-JUN-1997.
XX
PF 18-DEC-1996; 96WO-DK00542.
XX
PR 20-DEC-1995; 95DK-0001449.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Bal Schmidt P, Kjeldsen T, Pettersson A;
XX
DR WPI; 1997-341697/31.
DR P-PSDB; AAW19242.

XX
PT DNA construct encoding heterologous protein with signal and leader
PT peptide - attached by specific hexapeptide removable in vitro by
PT proteolysis, used to transform yeast host cells for protein
PT production

XX
PS Claim 8; Page 22; 32pp; English.
XX
CC A DNA sequence (AAT69680) in pAK749 codes for a polypeptide (AAW19242)
CC comprising the yeast YAP3 signal peptide and LA19 leader (see also
CC AAW19135), an extension hexapeptide (see also AAW19140), and the MI5
CC insulin precursor B chain(1-29)-Ser-Asp-Ala-Lys-A chain (1-21).
CC The construct was designed to improve expression yields of MI5
CC MI5 insulin in transformed yeast host cells. The MI5 insulin is
CC expressed as an N-terminally extended protein. The N-terminal
CC extension protects the MI5 insulin from proteolytic digestion
CC during fermentation and is then readily removed in vitro to provide
CC the mature polypeptide.

XX
SQ Sequence 550 BP; 177 A; 118 C; 94 G; 161 T; 0 other;
Query Match 49.8%; Score 91.2; DB 18; Length 550;
Best Local Similarity 73.9%; Pred. No. 2.3e-20;
Matches 133; Conservative 0; Mismatches 38; Indels 9; Gaps 1;
QY 1 ATGTTTCGTTAATCAGCACCTGTGGCGGCTCTCACCTGTGAGAGCTCTGTACCTGGTTGC 60
Db 322 AAGTTTCGTTAATCAGCACCTGTGGCGGTTCCACCTGGTTGAAGCTTTGTACTTGGTTGC 381
QY 61 GGTCAGCTGGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGGCATC 120
Db 382 GGTGAAAGAGGTTTCTTCTACACTCCTAAAGTCTGACGATGCTAAGGGT-----ATT 432
QY 121 GTTGACAACAATGCTGTACTAGCATCTGCTCTCTACAGCTGGAGAACTATTGTAACCTAG 180
Db 433 GTCGAGCAATGCTGTACTCTCCATCTGCTCTTGATCCAACTGGAAACTACTGCAACTAG 492

RESULT 11
AAQ86427
ID AAQ86427 standard; cDNA; 523 BP.
XX
AC AAQ86427;
XX
DT 20-NOV-1995 (first entry)
XX
DE Mating factor alpha 1-Insulin precursor MI5.
XX
KW Human insulin precursor MI5; diabetes; Zinc ion complex;
KW mating factor alpha 1; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 80..502
FT /*tag= a
XX
PN WO9507931-A.
XX
PD 23-MAR-1995.
XX
PF 16-SEP-1994; 94WO-DK00347.
XX
PR 17-SEP-1993; 93DK-0001044.
PR 02-FEB-1994; 94US-0190829.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Andersen AS, Halstrom JB, Havelund S, Jonassen I;
PI Markussen J;
XX
DR WPI; 1995-131314/17.
DR P-PSDB; AAR71693.
XX
PT Acylated insulin deriv. which may be present as a Zinc ion
PT complex - is used to treat diabetes and is rapid acting.
XX
PS Example 5; Pages 80-81; 100pp; English.
XX
CC AAQ86427 encodes AAR71693 mating factor alpha 1-Insulin precursor
CC MI5. MI5 comprises the B and A chains of a claimed human insulin
CC derivative, with a SDAK bridge connecting the B29 and A1 amino
CC acid residues. In the final claimed compsn. they are covalently
CC connected via disulphide bonds between Cys residues A7/B7 and
CC A20/B19. The derivative, which may be present as a zinc ion
CC complex, can be used as a fast action treatment for diabetes.
XX
SQ Sequence 523 BP; 156 A; 113 C; 102 G; 152 T; 0 other;
Query Match 49.1%; Score 89.8; DB 16; Length 523;
Best Local Similarity 74.0%; Pred. No. 6.7e-20;

SQ Sequence 415 BP; 117 A; 89 C; 86 G; 123 T; 0 other;

Query Match 50.8%; Score 93; DB 16; Length 415;
Best Local Similarity 75.1%; Pred. No. 5.3e-21;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;

QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTCTACCTGGTTGCGGT 63
DB 227 TTCGTTAACCAACACTTGTGGGTTCTCAGTTGGTTGAAGCTTTGACTTGGTTTGTGGT 286
QY 64 GAACGTGGTTTTTTCTACACCCCGGAAACCGGTGGTGGTCCGGGTAACGTTGGCATCGTT 123
DB 287 GAAAGAGGTTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGT-----ATCGTT 337
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAACCTAG 180
DB 338 GAACAATGCTGTACTTCTCTGTTCTTTGTACCAATTTGGAACAACTACTGTAACTAG 394

RESULT 8
AAQ86421
ID AAQ86421 standard; cDNA; 523 BP.
XX
AC AAQ86421;

XX 20-NOV-1995 (first entry)

XX Mating factor alpha 1-Insulin precursor M15.

XX Human insulin precursor M15; diabetes; Zinc ion complex;
KW mating factor alpha 1; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH 80..502
FT CDS /*tag= a

XX WO9507931-A.

XX 23-MAR-1995.

XX 16-SEP-1994; 94WO-DK00347.

XX 17-SEP-1993; 93DK-0001044.

XX 02-FEB-1994; 94US-0190829.

XX (NOVO) NOVO-NORDISK AS.

XX Andersen AS, Halstrom JB, Havelund S, Jonassen I;
PI Markussen J;

XX WPI; 1995-131314/17.

XX P-PSDB; AAR71690.

XX Acylated insulin deriv. which may be present as a Zinc ion
PT complex - is used to treat diabetes and is rapid acting.

XX Example 5; Page 74; 100pp; English.

XX AAQ86421 encodes AAR71690 mating factor alpha 1-Insulin precursor
CC M15. M15 comprises the B and A chains of a claimed human insulin
CC derivative, with a SDDAK bridge connecting the B29 and A1 amino
CC acid residues. In the final claimed compsn. they are covalently
CC connected via disulphide bonds between Cys residues A7/B7 and
CC A20/B19. The derivative, which may be present as a zinc ion
CC complex, can be used as a fast action treatment for diabetes.

XX Sequence 523 BP; 158 A; 104 C; 101 G; 160 T; 0 other;

Query Match 50.8%; Score 93; DB 16; Length 523;
Best Local Similarity 75.1%; Pred. No. 5.8e-21;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;

QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTCTACCTGGTTGCGGT 63
DB 335 TTCGTTAACCAACACTTGTGGGTTCTCAGTTGGTTGAAGCTTTGACTTGGTTTGTGGT 394
QY 64 GAACGTGGTTTTTTCTACACCCCGGAAACCGGTGGTGGTCCGGGTAACGTTGGCATCGTT 123
DB 395 GAAAGAGGTTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGT-----ATCGTT 445
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAACCTAG 180
DB 446 GAACAATGCTGTACTTCTCTGTTCTTTGTACCAATTTGGAACAACTACTGTAACTAG 502

RESULT 9
AAT37134
ID AAT37134 standard; DNA; 258 BP.
XX
AC AAT37134;

XX 15-APR-1997 (first entry)

XX Proinsulin coding sequence.

XX Proinsulin; human; beta-turn peptide; Insulin; A chain; protein folding;
KW enzymatic cleavage; B chain; miniproinsulin; enzymatic hydrolysis; ss.
XX Synthetic.

XX GB2298206-A.

XX 28-AUG-1996.

XX 14-FEB-1996; 96GB-0002998.

XX 15-FEB-1995; 95KR-0002751.

XX (HANI-) HANIL SYNTHETIC FIBER CO LTD.

XX Chang S, Kim C, Kim D, Shin H;

XX WPI; 1996-373460/38.

XX New human pro-insulin derivs. - comprising A and B chains linked by
PT small beta-turn peptide

XX Example 4; Page 18; 52pp; English.

XX This sequence represents a synthetic human proinsulin coding sequence.
CC The codon preference of this sequence has been altered to favour the
CC preferred codons in E. coli. This sequence can be used in the human
CC proinsulin derivatives of the invention. In the derivatives of the
CC invention, a beta-turn peptide (see AAW03935-W03938) is used to join the
CC human insulin A and B chains. Beta-turn peptides are thought to be
CC sites for initiation of protein folding. This is thought to be due to
CC the fact that they are determined by short-range interactions, and
CC therefore limit the conformational space available to the polypeptide
CC chain. Beta-turns also play a valuable role in relation to enzymatic
CC cleavage. By using the insulin derivatives (also referred to as
CC miniproinsulin), refolding and hydrolysing processes can be carried out
CC more efficiently than with proinsulin (which contains a bulky C-peptide).
CC The derivatives can be recovered from transformed host cells in higher
CC refolding yields than human proinsulin. The derivatives are also readily
CC converted to insulin by enzymatic hydrolysis, such as with trypsin and
CC carboxypeptidase B. The insulin derivatives are easier to produce than
CC current derivatives. By using the derivatives of the invention, insulin
CC can be produced at higher yields than with current techniques.

XX Sequence 258 BP; 50 A; 67 C; 75 G; 66 T; 0 other;

Query Match 50.4%; Score 92.2; DB 17; Length 258;
Best Local Similarity 96.9%; Pred. No. 8.4e-21;
Matches 94; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```
Query Match          48.3%; Score 88.4; DB 17; Length 351;
Best Local Similarity 75.0%; Pred. No. 1.7e-19;
Matches 132; Conservative 0; Mismatches 26; Indels 18; Gaps 1;

Qy 4 TTCGTTAATCAGCACCTGTGCGGCTCTCACCTGTGTAAGACTCTGTACCTGGTTTGGCGT 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 TTTGTCACACGACCTGTGTTCTTCACCTGTGTTGAAGCACTGTACCTGGTATGTGGC 252

Qy 64 GAACGTGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTTGCGATCGTT 123
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 GAACGTGTTTTTCTACACTCTCTAAAACC-----CGCGGCATCGTT 294

Qy 124 GAACAATGCTGTACTAGCATCTCTCTCTACACGCTGGAGAACTATTGTAACTA 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 295 GAACAGTGTCTACTCTATCTATCTGTCTCCCTGTACCAACTGGAGAACTACTGCAATTA 350

RESULT 14
AAQ86412
ID AAQ86412 standard; cDNA; 415 BP.
XX
AC AAQ86412;
XX
DT 20-NOV-1995 (first entry)
XX
DE Yeast signal/leader Lac212spX3-Insulin precursor AlaA21, ThrB3.
XX
KW Human insulin precursor; diabetes; Zinc ion complex;
KW synthetic yeast signal/leader; Lac212spX3; AlaA21, ThrB3; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 80..394
FT /*tag= a
XX
PN WO9507931-A.
XX
PD 23-MAR-1995.
XX
PF 16-SEP-1994; 94WO-DK00347.
XX
PR 17-SEP-1993; 93DK-0001044.
PR 02-FEB-1994; 94US-0190829.
XX
PA (NOVO ) NOVO-NORDISK AS.
XX
PI Andersen AS, Halstrom JB, Havelund S, Jonassen I;
PI Markussen J;
XX
DR WPI: 1995-131314/17.
DR P-PSDB; AAR71687.
XX
PT Acylated insulin deriv. which may be present as a Zinc ion
PT complex - is used to treat diabetes and is rapid acting.
XX
PS Example 2; Page 68; 100pp; English.
XX
CC AAQ86412 encodes AAR71687 the synthetic yeast signal/leader
CC Lac212spX3-Insulin precursor AlaA21, ThrB3. AlaA21,
CC ThrB3 comprises the B and A chains of a claimed human insulin
CC derivative, with a SDDAK bridge connecting the B29 and A1 amino
CC acid residues. In the final claimed compsn. they are covalently
CC connected via disulphide bonds between Cys residues A7/B7 and
CC A20/B19. The derivative, which may be present as a zinc ion
CC complex, can be used as a fast action treatment for diabetes.
XX
SQ Sequence 415 BP; 114 A; 89 C; 87 G; 125 T; 0 other;

Query Match          48.2%; Score 88.2; DB 16; Length 415;
Best Local Similarity 73.4%; Pred. No. 2.1e-19;
Matches 130; Conservative 0; Mismatches 38; Indels 9; Gaps 1;

Qy 4 TTCGTTAATCAGCACCTGTGCGGCTCTCACCTGTGTAAGACTCTGTACCTGGTTTGGCGT 63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 TTCGTTACTCAACACTTGTGCGGTCTCTCACTTGGTTGAAGCTTTGTACTTGGTTTGGT 286
```



```
Qy 64 GAACGTGGTTTTTTCTACACCCCGAACCGTGGTCCGGGTAAACGTGGCATCGTT 123
    ||| | ||||| ||||| | | | | | |||
Db 287 GAAAGAGGTTTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGT-----ATCGTT 337
    ||||| ||||| ||||| | | | | | |||
Qy 124 GAACAATGCTGTACTAGCATCTGCTCTCTACCAAGCTGGAGAACTATTGTAACTAG 180
    ||||| ||||| ||||| | | | | | |||
Db 338 GAACAATGTTGTACTTCTATCTGTCTTTGTACCAATTGGAAAACTACTGTGGTTAG 394
    ||||| ||||| ||||| | | | | | |||
```

Search completed: February 4, 2003, 06:59:04
Job time : 275 secs

THIS PAGE BLANK (USPTO)

QY 4 TTCGTTAATCAGCACCTGTGGGGTCTCACTGCTAGAACGCTGTGTACCTGGTTTGGCGT 63
Db 227 TTCGTTAACCAACACTGTGGGGTCTCACTTGGTTGAAGCTTTGTACTTGGTTTGGT 286
QY 64 GAACGTTGTTTTCTACACCCGAAACGGTGGTGGTCCGGGTAACGCTGGCATCGTT 123
Db 287 GAAAGAGTTTCTTCTACACTCCAAAGCTGACACGCTAAGGT-----ATCGTT 337
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTCTACAGCTGGAGAACTATTGTAACTAG 180
Db 338 GAACAATGTTGTACTTCTATCTGTTCTTTGTACCAATTGGAARACTACTGTAACTAG 394
RESULT 2
US-08-975-365-14
: Sequence 14, Application US/08975365
: Patent No. 6011007
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: APPLICANT: Markussen, Jan
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/975,365
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/400,256
: FILING DATE: 03-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985.220-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 415 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 80..391
US-08-975-365-14
Query Match 50.8%; Score 93; DB 3; Length 415;
Best Local Similarity 75.1%; Pred. No. 2.9e-22;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;
QY 4 TTCGTTAATCAGCACCTGTGGGGTCTCACTGCTAGAACGCTGTGTACCTGGTTTGGCGT 63
Db 227 TTCGTTAACCAACACTGTGGGGTCTCACTTGGTTGAAGCTTTGTACTTGGTTTGGT 286
QY 64 GAACGTTGTTTTCTACACCCGAAACGGTGGTGGTCCGGGTAACGCTGGCATCGTT 123

Db 287 GAAAGAGTTTCTTCTACACTCCAAAGCTGACACGCTAAGGT-----ATCGTT 337
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTCTACAGCTGGAGAACTATTGTAACTAG 180
Db 338 GAACAATGTTGTACTTCTATCTGTTCTTTGTACCAATTGGAARACTACTGTAACTAG 394
RESULT 3
US-08-400-256-32
: Sequence 32, Application US/08400256
: Patent No. 5750497
: GENERAL INFORMATION:
: APPLICANT: Havelund, Svend
: APPLICANT: Halstrom, John
: APPLICANT: Jonassen, Ib
: APPLICANT: Andersen, Asger Sloth
: APPLICANT: Markussen, Jan
: TITLE OF INVENTION: ACYLATED INSULIN
: NUMBER OF SEQUENCES: 49
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
: STREET: 405 Lexington Avenue, 64th Floor
: CITY: New York
: STATE: New York
: COUNTRY: United States of America
: ZIP: 10174-6401
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/400,256
: FILING DATE: 03-MAR-1995
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Lambiris, Elias J.
: REGISTRATION NUMBER: 33,728
: REFERENCE/DOCKET NUMBER: 3985.220-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 32:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 523 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 80..499
US-08-400-256-32
Query Match 50.8%; Score 93; DB 1; Length 523;
Best Local Similarity 75.1%; Pred. No. 3.1e-22;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;
QY 4 TTCGTTAATCAGCACCTGTGGGGTCTCACTGCTAGAACGCTGTGTACCTGGTTTGGCGT 63
Db 335 TTCGTTAACCAACACTGTGGGGTCTTCTCACTTGGTTGAAGCTTTGTACTTGGTTTGGT 394
QY 64 GAACGTTGTTTTCTACACCCGAAACGGTGGTGGTCCGGGTAACGCTGGCATCGTT 123
Db 395 GAAAGAGTTTCTTCTACACTCCAAAGCTGACACGCTAAGGT-----ATCGTT 445
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTCTACAGCTGGAGAACTATTGTAACTAG 180
Db 446 GAACAATGTTGTACTTCTATCTGTTCTTTGTACCAATTGGAARACTACTGTAACTAG 502
RESULT 4

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 06:45:23 ; Search time 55 seconds
(without alignments)
1020.397 Million cell updates/sec

Title: US-09-706-690-3

Perfect score: 183

Sequence: 1 atgttcgttaatcagcacct.....agaactatgttaactagtaa 183

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCRTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 93 | 50.8 | 415 | 1 | US-08-400-256-14 |
| 2 | 93 | 50.8 | 415 | 3 | US-08-975-365-14 |
| 3 | 93 | 50.8 | 523 | 1 | US-08-400-256-32 |
| 4 | 93 | 50.8 | 523 | 3 | US-08-975-365-32 |
| 5 | 92.2 | 50.4 | 258 | 2 | US-08-600-783-17 |
| 6 | 89.8 | 49.1 | 523 | 1 | US-08-400-256-41 |
| 7 | 89.8 | 49.1 | 523 | 3 | US-08-975-365-41 |
| 8 | 89 | 48.6 | 192 | 1 | US-08-030-731A-45 |
| 9 | 89 | 48.6 | 192 | 3 | US-08-975-365-45 |
| 10 | 88.4 | 48.3 | 352 | 1 | US-08-030-731A-46 |
| 11 | 88.2 | 48.2 | 415 | 1 | US-08-967-867-6 |
| 12 | 88.2 | 48.2 | 415 | 3 | US-08-400-256-23 |
| 13 | 88.2 | 48.2 | 415 | 1 | US-08-400-256-29 |
| 14 | 88.2 | 48.2 | 415 | 3 | US-08-975-365-23 |
| 15 | 88 | 48.1 | 354 | 3 | US-08-975-365-29 |
| 16 | 86.8 | 47.4 | 523 | 1 | US-08-967-867-5 |
| 17 | 86.8 | 47.4 | 523 | 3 | US-08-400-256-17 |
| 18 | 86.6 | 47.3 | 415 | 1 | US-08-975-365-17 |
| 19 | 86.6 | 47.3 | 415 | 3 | US-08-400-256-20 |
| 20 | 86.6 | 47.3 | 415 | 1 | US-08-400-256-26 |
| 21 | 86.6 | 47.3 | 415 | 3 | US-08-975-365-20 |
| 22 | 84.8 | 46.3 | 162 | 4 | US-08-975-365-26 |
| 23 | 84 | 45.9 | 219 | 1 | US-08-981-988A-43 |
| 24 | 84 | 45.9 | 219 | 1 | US-08-468-674B-70 |
| 25 | 84 | 45.9 | 550 | 4 | US-08-780-571-70 |
| 26 | 82.6 | 45.1 | 297 | 1 | US-09-012-669F-1 |
| 27 | 82.6 | 45.1 | 297 | 1 | US-08-468-674B-44 |
| | | | | | Sequence 44, Appl |

| | | | | | | |
|------|------|------|-----|---|-------------------|--------------------|
| 28 | 82.6 | 45.1 | 409 | 1 | US-08-400-256-35 | Sequence 35, Appl |
| 29 | 82.6 | 45.1 | 409 | 3 | US-08-975-365-35 | Sequence 35, Appl |
| 30 | 82.6 | 45.1 | 476 | 1 | US-08-446-846-2 | Sequence 2, Appl |
| 31 | 82.6 | 45.1 | 500 | 4 | US-09-012-669F-2 | Sequence 2, Appl |
| 32 | 82.6 | 45.1 | 594 | 1 | US-08-932-082-18 | Sequence 18, Appl |
| 33 | 81.4 | 44.5 | 511 | 1 | US-08-400-256-38 | Sequence 38, Appl |
| 34 | 81.4 | 44.5 | 511 | 3 | US-08-975-365-38 | Sequence 38, Appl |
| 35 | 81.4 | 44.5 | 535 | 1 | US-08-400-256-44 | Sequence 44, Appl |
| 36 | 81.4 | 44.5 | 535 | 3 | US-08-975-365-44 | Sequence 44, Appl |
| 37 | 81.4 | 44.5 | 538 | 1 | US-08-400-256-47 | Sequence 47, Appl |
| 38 | 81.4 | 44.5 | 538 | 3 | US-08-975-365-47 | Sequence 47, Appl |
| 39 | 81 | 44.3 | 372 | 1 | US-08-468-674B-40 | Sequence 40, Appl |
| 40 | 81 | 44.3 | 372 | 1 | US-08-780-571-40 | Sequence 40, Appl |
| 41 | 77.8 | 42.5 | 281 | 1 | US-07-764-655D-12 | Sequence 12, Appl |
| 42 | 77.8 | 42.5 | 281 | 1 | US-07-764-655D-13 | Sequence 13, Appl |
| C 42 | 77.8 | 42.5 | 281 | 6 | 5514646-1 | Patent No. 5514646 |
| 44 | 74.4 | 40.7 | 115 | 1 | US-07-764-655D-16 | Sequence 16, Appl |
| C 45 | 74.4 | 40.7 | 116 | 1 | US-07-764-655D-17 | Sequence 17, Appl |

ALIGNMENTS

RESULT 1

US-08-400-256-14

; Sequence 14, Application US/08400256

; Patent No. 5750497

; GENERAL INFORMATION:

; APPLICANT: Havelund, Svend

; APPLICANT: Halstrom, John

; APPLICANT: Jonassen, Ib

; APPLICANT: Andersen, Asger Sloth

; APPLICANT: Markussen, Jan

; TITLE OF INVENTION: ACYLATED INSULIN

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.

; STREET: 405 Lexington Avenue, 64th Floor

; CITY: New York

; STATE: New York

; COUNTRY: United States of America

; ZIP: 10174-6401

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/400,256

; FILING DATE: 03-MAR-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Lambiris, Elias J.

; REGISTRATION NUMBER: 33,728

; REFERENCE/DOCKET NUMBER: 3985.220-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123

; TELEFAX: 212-878-9655

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 415 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 80..391

; US-08-400-256-14

Query Match 50.8%; Score 93; DB 1; Length 415;
Best Local Similarity 75.1%; Pred. No. 2.9e-22;
Matches 133; Conservative 0; Mismatches 35; Indels 9; Gaps 1;

RESULT 6
US-08-400-256-41
; Sequence 41, Application US/08400256
; Patent No. 5750497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; NAME/KEY: CDS
; LOCATION: 80..499
US-08-400-256-41
Query Match 49.1%; Score 89.8; DB 1; Length 523;
Best Local Similarity 74.0%; Pred. No. 3.7e-21;
Matches 131; Conservative 0; Mismatches 37; Indels 9; Gaps 1;
QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGTGTAGAGCTCTGTACCTGTGTTGGGT 63
Db 335 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGTGTAGAGCTCTGTACCTGTGTTGGGT 394
QY 64 GAACGTGCTTTTCTACACCCGAAACCGGTGGTCCGGGTAAACGTTGTCATCGTT 123
Db 395 GAAGAGGTTTCTTCTACACTCTTAAGTCTGACGATGCTAAGGT-----ATTGTC 445
QY 124 GAACAATGCTGTAGATCTGTCTCTTACAGCTGGAGAACTATTTGTAAC TAG 180
Db 446 GAGCAATGCTGTACCTCCATCTGCTCTTGTACCAATTGGAAAACTACTGCAACTAG 502
RESULT 7
US-08-975-365-41
; Sequence 41, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John

; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/975,365
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; NAME/KEY: CDS
; LOCATION: 80..499
US-08-975-365-41
Query Match 49.1%; Score 89.8; DB 3; Length 523;
Best Local Similarity 74.0%; Pred. No. 3.7e-21;
Matches 131; Conservative 0; Mismatches 37; Indels 9; Gaps 1;
QY 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGTGTAGAGCTCTGTACCTGTGTTGGGT 63
Db 335 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGTGTAGAGCTCTGTACCTGTGTTGGGT 394
QY 64 GAACGTGCTTTTCTACACCCGAAACCGGTGGTCCGGGTAAACGTTGTCATCGTT 123
Db 395 GAAGAGGTTTCTTCTACACTCTTAAGTCTGACGATGCTAAGGT-----ATTGTC 445
QY 124 GAACAATGCTGTAGATCTGTCTCTTACAGCTGGAGAACTATTTGTAAC TAG 180
Db 446 GAGCAATGCTGTACCTCCATCTGCTCTTGTACCAATTGGAAAACTACTGCAACTAG 502
RESULT 8
US-08-030-731A-45
; Sequence 45, Application US/08030731A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Koller, Klaus-Peter
; APPLICANT: Riess, Guenther Johannes
; APPLICANT: Uhlmann, Eugen
; APPLICANT: Wallmeier, Holger
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; TITLE OF INVENTION: Proteins in Streptomyces
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:

US-08-975-365-32
; Sequence 32, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 80...499
US-08-975-365-32

Query Match 50.8%; Score 93; DB 3; Length 523;
Best Local Similarity 75.1%; Pred. No. 3.le-22;
Matches 133; Conservative 0; Mismatches 55; Indels 9; Gaps 1;
QY 4 TTCGTTAATCAGCACCTCTGCGGCTCTCACCTGGTAGAAGCTCTGTACTGTTGGCGT 63
Db 335 TTCGTTAACCAACACTCTGCGGTTCTGCTTGTGTAAGCTTGTACTTGGTTGGT 394
QY 64 GAACGTGGTTTTTCTACACCCCGAACCAGGTGGTGGTGGGTAACGTGGCATCGTT 123
Db 395 GAAGAGGTTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGT-----ATCGTT 445
QY 124 GAACAATGCTGTACTAGCATCTGCTCTCTACACGCTGGAGAACTATTGTAACCTAG 180
Db 446 GAACANTGTTGACTTCTATCTGTCTTTGTACCAATGGAAACTACTGTAACTAG 502

RESULT 5
US-08-600-783-17
; Sequence 17, Application US/08600783
; Patent No. 5962267
; GENERAL INFORMATION:
; APPLICANT: SHIN, Hang Cheol
; APPLICANT: CHANG, Seung Gu

; APPLICANT: KIM, Dae Young
; APPLICANT: KIM, Chong Suh
; TITLE OF INVENTION: Proinsulin Derivative and Process
; TITLE OF INVENTION: for Producing Human Insulin
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SHIN, Hang Cheol
; STREET: Ssangma-Hanshin Apt. 102-1206,
; CITY: Kwangmyung-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 423-030
; ADDRESSEE: CHANG, Seung Gu
; STREET: Hyundai Apt. 71-203, Apjukong-dong,
; STREET: Kangnam-ku
; CITY: Seoul
; STATE: Seoul
; COUNTRY: Republic of Korea
; ZIP: 135-110
; ADDRESSEE: KIM, Dae Young
; STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
; STREET: Sosa-ku
; CITY: Bucheon-shi
; STATE: Kyungki-do
; COUNTRY: Republic of Korea
; ZIP: 422-230
; ADDRESSEE: KIM, Chong Suh
; STREET: Garden Heights Apt. 202-801, #100,
; STREET: Hwangkeum-dong, Soosung-ku
; CITY: Taegu
; STATE: Taegu
; COUNTRY: Republic of Korea
; ZIP: 706-040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage
; COMPUTER: IBM PC/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,783
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: KR 95-2751
; FILING DATE: 15-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahan Islam
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-1000
; TELEFAX: (212) 953-7249
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 258 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-600-783-17

Query Match 50.4%; Score 92.2; DB 2; Length 258;
Best Local Similarity 96.9%; Pred. No. 4.5e-22;
Matches 94; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 4 TTCGTTAATCAGCACCTCTGCGGCTCTCACCTGGTAGAAGCTCTGTACTGTTGGCGT 63
Db 1 TTCGTTAATCAGCACCTCTGCGGCTCTCACCTGGTAGAAGCTCTGTACTGTTGGCGT 60
QY 64 GAACGTGGTTTTTCTACACCCCGAACCAGGTGGT 100
Db 61 GAACGTGGTTTTTCTACACCCCGAACCAGGTGGT 97

; LOCATION: 5..192
; OTHER INFORMATION: /note= "Sequence ID No. 5426036 46 is
; complementary to Sequence ID No. 5426036 45 at positions
; OTHER INFORMATION: 5-192 of Sequence ID No. 5426036 45."
US-08-030-731A-46

Query Match 48.6%; Score 89; DB 1; Length 192;
Best Local Similarity 73.8%; Pred. No. 4.9e-21;
Matches 135; Conservative 0; Mismatches 30; Indels 18; Gaps 1;

Qy 1 ATGTTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTGAAGACTCTGTACCTGTTTGC 60

Db 184 AAGTTCGTCAACAGCACCTGTGGGCTCGCACCTGTGGAGCCCTCTACCTGTTGTC 125

Qy 61 GGTGAAGCTGTTTTTCTACACCCGAAACCGGTGGTCCGGGTAAACGTGGGCATC 120

Db 124 GGGGAGGCGGCTTCTTCTACACCCCAAGACC-----AAGGGCATC 83

Qy 121 GTTCAACAATCCTACTAGCATCTGCTCTCTACAGCTGGAGAACTATTGTAACATG 180

Db 82 GTGAGAGTGTGTACGTCCATCTGCTCCCTTACCAGCTCGAGAACTACTGCAACTAG 23

Qy 181 TAA 183

Db 22 TAA 20

RESULT 10
US-08-967-867-6
; Sequence 6, Application US/08967867
; Patent No. 6001604
; GENERAL INFORMATION:
; APPLICANT: HARTMAN, JACOB R.
; APPLICANT: MENDELOVITZ, SIMONA
; APPLICANT: GORECKI, MARIAN
; TITLE OF INVENTION: GENERATION OF HUMAN INSULIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM LLP
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,867
; FILING DATE: 12-NOV-1997
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,454
; FILING DATE: 29-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41425-A/JPW/GJG

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..352
US-08-967-867-6

Query Match 48.3%; Score 88.4; DB 3; Length 352;
Best Local Similarity 75.0%; Pred. No. 9.6e-21;
Matches 132; Conservative 0; Mismatches 26; Indels 18; Gaps 1;

Qy 4 TTCCTTAATCAGCACCTGTGGGCTCTCACCTGGTGAAGACTCTGTACCTGTTTGGGT 63

Db 193 TTTGTCAACAGCACCTGTGGTCTCTCACCTGGTGAAGCACTGTACCTGTTATGTGGC 252

Qy 64 GAAGCTGGTTTTTCTACACCCGAAACCGGTGGTCCGGGTAAACGTGGGCATCGTT 123

Db 253 GAAGCTGGTTTTTCTACACTCTCTAAAAACC-----CGCGGCATCGTT 294

Qy 124 GAACAATCCTCTACTAGCATCTGCTCTCTACAGCTGGAGAACTATTGTAACATA 179

Db 295 GAACAGTGTGTACCTCTATCTGTTCCTGTACCACTGGAGAACTACTGCAATTA 350

RESULT 11
US-08-400-256-23
; Sequence 23, Application US/08400256
; Patent No. 5750497
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan

; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 80..391
US-08-400-256-23

Query Match 48.2%; Score 88.2; DB 1; Length 415;
Best Local Similarity 73.4%; Pred. No. 1.2e-20;
Matches 130; Conservative 0; Mismatches 38; Indels 9; Gaps 1;

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/030,731A
FILING DATE: 12-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/189,840
FILING DATE: 03-MAY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/430,622
FILING DATE: 01-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/687,610
FILING DATE: 19-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/735,757
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 37 14 866.4
FILING DATE: 05-MAY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 38 37 273.8
FILING DATE: 03-NOV-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 39 27 449.7
FILING DATE: 19-AUG-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 40 12 818.0
FILING DATE: 21-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Kirschner Michael K.
REGISTRATION NUMBER: 34,851
REFERENCE/DOCKET NUMBER: 02481-0593-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 192 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-030-731A-45

Query Match 48.6%; Score 89; DB 1; Length 192;
Best Local Similarity 73.8%; Pred. No. 4.9e-21;
Matches 135; Conservative 0; Mismatches 30; Indels 18; Gaps 1;
Qy 1 ATGTTTCGTTAATCAGCACCTGTGGCGGCTCTCACCTGGTAGAAGCTCTGTACTGTGGTTTC 60
Db 13 AAGTTTCGTTCAACACACACCTGTGGCGGCTCGACCTGTGGAGGCCCTCTACTGTGGTGC 72
Qy 61 GGTGAACGTGGTTTTTCTACACCCCGAAACCCGGTGGTGGTCCGGGTAACGTTGGCATC 120
Db 73 GGGGAGCGCGGCTTCTTCTACACCCCGAAGACC-----AAGGGCATC 114
Qy 121 GTTGAACAACTGCTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAACATAG 180
Db 115 GTGGAGCAGTGTGTACGTCCATCTGCTCCCTCTACCAAGCTCGAGAACTACTGCAACTAG 174
Qy 181 TAA 193

Db 175 TAA 177
RESULT 9
US-08-030-731A-46/c
; Sequence 46, Application US/08030731A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Koller, Klaus-Peter
; APPLICANT: Riess, Guenther Johannes
; APPLICANT: Uhlmann, Eugen
; APPLICANT: Wallmeier, Holger
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,731A
; FILING DATE: 12-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/189,840
; FILING DATE: 03-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/430,622
; FILING DATE: 01-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/687,610
; FILING DATE: 19-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,757
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 37 14 866.4
; FILING DATE: 05-MAY-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 38 37 273.8
; FILING DATE: 03-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 39 27 449.7
; FILING DATE: 19-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 40 12 818.0
; FILING DATE: 21-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Kirschner Michael K.
; REGISTRATION NUMBER: 34,851
; REFERENCE/DOCKET NUMBER: 02481-0593-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature

; Sequence 29, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asger Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 415 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 80..391
US-08-975-365-29

Query Match 48.2%; Score 88.2; DB 3; Length 415;
Best Local Similarity 73.4%; Pred. No. 1.2e-20;
Matches 130; Conservative 0; Mismatches 38; Indels 9; Gaps 1;
QY 4 TTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGGGT 63
Db 227 TTGTTACTCAACACTTTGTGGGTTTCTACCTTTGTTGAAGCTTTGTTGTTGGT 286
QY 64 GAACGTGGTTTTTTTCTACACCCCGAAGCGGTGGTCCGGTAAACCTGGCATCGTT 123
Db 287 GAAGAGGTTTCTTCTACACTCCAAAGCTGTGACGACGCTAAGGT-----ATCGTT 337
QY 124 GAACAATGCTGTACTAGCATCTCTCTCTACAGCTGGAGAACATTTGTAACCTAG 180
Db 338 GAACAATGTTGTACTTCTCTCTCTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTAG 394

RESULT 15

US-08-967-867-5
; Sequence 5, Application US/08967867
; Patent No. 6001604
; GENERAL INFORMATION:
; APPLICANT: HARTMAN, JACOB R.
; APPLICANT: MENDELOVITZ, SIMONA
; APPLICANT: GORECKI, MARIAN

; TITLE OF INVENTION: GENERATION OF HUMAN INSULIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM LLP
; STREET: 1185 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,867
; FILING DATE: 12-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,454
; FILING DATE: 29-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41425-A/JPW/GJJ
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 354 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..354
US-08-967-867-5

Query Match 48.1%; Score 88; DB 3; Length 354;
Best Local Similarity 74.4%; Pred. No. 1.3e-20;
Matches 131; Conservative 0; Mismatches 30; Indels 15; Gaps 1;
QY 4 TTGCTTAATCAGCACCTGTGGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGGGT 63
Db 193 TTGTCACACGACGACCTGTGTGGTCTCACCTAATTGAAGCACTGTACCTGGTAGTGGC 252
QY 64 GAACGTGGTTTTTTTCTACACCCCGAAGCGGTGGTCCGGTAAACCTGGCATCGTT 123
Db 253 GAACGTGGTTTCTTCTACACTCCCTAAAC-----AAAGCGGGCATCGTT 297
QY 124 GAACAATGCTGTACTAGCATCTCTCTCTACAGCTGGAGAACATTTGTAACCTA 179
Db 298 GAACAGTGTGTACCTCTATCTGTTCCTGTACCAACTGGAGAACTACTGCAATTA 353

Search completed: February 4, 2003, 07:00:16
Job time : 61 secs

| | | | |
|---|-----|--|-----|
| Qy | 4 | TTCGTTAATCAGACCTGTGGCGGTCTCTACCTGGTAGAAGCTCTGTACTCTGGTTTGGCGGT | 63 |
| Db | 227 | TTCGTTTACTCAACACTTTGTGGCGGTCTCTACCTGGTTGAAGCTTTGTACTTGGTTTGGT | 286 |
| Qy | 64 | GAACGTGGTTTTTCTACACCCCGAAAAACCGTGGTGGTCCGGGTAAACGTGGCATCGTT | 123 |
| Db | 287 | GAAGAAGGTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGGT-----ATCGTT | 337 |
| Qy | 124 | GAACAATGCTCTACTAGCATCTGCTCTCTCTACGAGCTGGAGAACTATTGTAACCTAG | 180 |
| Db | 338 | GAACAATGTTGCTACTTCTATCTGTCTTGTGACCAATTGGAAACTACTGTGCTTAG | 394 |
| <p>RESULT 12</p> <p>US-08-400-256-29</p> <p>Sequence 29, Application US/08400256</p> <p>Patent No. 5750497</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Havelund, Svend</p> <p>APPLICANT: Halstrom, John</p> <p>APPLICANT: Jonassen, Ib</p> <p>APPLICANT: Andersen, Asser Sloth</p> <p>APPLICANT: Markussen, Jan</p> <p>TITLE OF INVENTION: ACYLATED INSULIN</p> <p>NUMBER OF SEQUENCES: 49</p> <p>CORRESPONDENCE ADDRESS: 49</p> <p>ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.</p> <p>STREET: 405 Lexington Avenue, 64th Floor</p> <p>CITY: New York</p> <p>STATE: New York</p> <p>COUNTRY: United States of America</p> <p>ZIP: 10174-6401</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patentin Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/400,256</p> <p>FILING DATE: 03-MAR-1995</p> <p>CLASSIFICATION: 514</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Lambiris, Elias J.</p> <p>REGISTRATION NUMBER: 33,728</p> <p>REFERENCE/DOCKET NUMBER: 3985.220-US</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 212-867-0123</p> <p>TELEFAX: 212-878-9655</p> <p>INFORMATION FOR SEQ ID NO: 29:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 415 base pairs</p> <p>TYPE: nucleic acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: cDNA</p> <p>FEATURE:</p> <p>NAME/KEY: CDS</p> <p>LOCATION: 80..391</p> <p>US-08-400-256-29</p> | | | |
| <p>Query Match 48.2%; Score 88.2; DB 1; Length 415;</p> <p>Best Local Similarity 73.4%; Pred. No. 1.2e-20;</p> <p>Matches 130; Conservative 0; Mismatches 38; Indels 9; Gaps 1;</p> | | | |
| Qy | 4 | TTCGTTAATCAGACCTGTGGCGGTCTCTACCTGGTAGAAGCTCTGTACTCTGGTTTGGCGGT | 63 |
| Db | 227 | TTCGTTTACTCAACACTTTGTGGCGGTCTCTACCTGGTTGAAGCTTTGTACTTGGTTTGGT | 286 |
| Qy | 64 | GAACGTGGTTTTTCTACACCCCGAAAAACCGTGGTGGTCCGGGTAAACGTGGCATCGTT | 123 |
| Db | 287 | GAAGAAGGTTCTTCTACACTCCAAAGTCTGACGACGCTAAGGGT-----ATCGTT | 337 |
| Qy | 124 | GAACAATGCTCTACTAGCATCTGCTCTCTCTACGAGCTGGAGAACTATTGTAACCTAG | 180 |

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 06:46:27 ; Search time 62 seconds
(without alignments)
1399.073 Million cell updates/sec

Title: US-09-706-690-3
Perfect score: 183
Sequence: 1 atgttcgtaatacagcacct.....agaactattgtaactagtaa 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues
Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PublishedApplications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 85.6 | 46.8 | 550 | 10 | US-09-740-359-11 |
| 2 | 85.6 | 46.8 | 550 | 10 | US-09-894-711-11 |
| 3 | 84 | 45.9 | 550 | 10 | US-09-894-711-17 |
| 4 | 82.6 | 45.1 | 594 | 9 | US-09-861-687-18 |
| 5 | 82.4 | 45.0 | 550 | 10 | US-09-740-359-9 |
| 6 | 82.4 | 45.0 | 550 | 10 | US-09-894-711-9 |
| 7 | 80.8 | 44.2 | 550 | 10 | US-09-736-611-11 |
| 8 | 80.8 | 44.2 | 600 | 10 | US-09-740-359-6 |
| 9 | 80.8 | 44.2 | 600 | 10 | US-09-894-711-6 |
| 10 | 79.2 | 43.3 | 600 | 10 | US-09-736-611-7 |
| 11 | 77.6 | 42.4 | 550 | 10 | US-09-736-611-5 |
| 12 | 72.4 | 39.6 | 600 | 10 | US-09-740-359-4 |
| 13 | 72.4 | 39.6 | 600 | 10 | US-09-894-711-4 |
| 14 | 72.4 | 39.6 | 600 | 10 | US-09-280-030-49 |
| 15 | 66.4 | 36.3 | 351 | 10 | US-09-280-030-48 |
| 16 | 66.4 | 36.3 | 390 | 10 | US-09-280-030-48 |
| 17 | 65 | 35.5 | 450 | 9 | US-09-804-409A-10 |
| 18 | 65 | 35.5 | 498 | 10 | US-09-925-297-346 |
| 19 | 65 | 35.5 | 1182 | 10 | US-09-919-344-27 |

| | | | | | | |
|------|------|------|------|----|--------------------|-------------------|
| 20 | 62.8 | 34.3 | 598 | 9 | US-10-013-032-11 | Sequence 11, Appl |
| 21 | 62.8 | 34.3 | 633 | 9 | US-10-013-032-12 | Sequence 12, Appl |
| 22 | 62.8 | 34.3 | 666 | 9 | US-10-013-032-13 | Sequence 13, Appl |
| 23 | 58.6 | 32.0 | 390 | 10 | US-09-833-381-939 | Sequence 939, App |
| c 24 | 51.2 | 28.0 | 574 | 9 | US-10-025-380-918 | Sequence 918, App |
| c 25 | 51.2 | 28.0 | 574 | 10 | US-09-922-217-918 | Sequence 918, App |
| c 26 | 51.2 | 28.0 | 574 | 10 | US-09-833-263-918 | Sequence 918, App |
| c 27 | 51.2 | 28.0 | 1356 | 10 | US-09-962-436-293 | Sequence 293, App |
| 28 | 51.2 | 28.0 | 1356 | 10 | US-09-954-456-294 | Sequence 294, App |
| 29 | 51.2 | 28.0 | 1356 | 10 | US-09-880-107-2092 | Sequence 2092, Ap |
| 30 | 51.2 | 28.0 | 4350 | 9 | US-10-125-181-37 | Sequence 37, Appl |
| c 31 | 49.8 | 27.2 | 447 | 9 | US-10-025-380-917 | Sequence 917, App |
| c 32 | 49.8 | 27.2 | 447 | 10 | US-09-922-217-917 | Sequence 917, App |
| c 33 | 49.8 | 27.2 | 447 | 10 | US-09-833-263-917 | Sequence 917, App |
| c 34 | 47.8 | 26.1 | 107 | 10 | US-09-736-611-14 | Sequence 14, Appl |
| 35 | 46.2 | 25.2 | 79 | 9 | US-09-861-687-12 | Sequence 12, Appl |
| c 36 | 44.6 | 24.4 | 82 | 9 | US-09-861-687-26 | Sequence 26, Appl |
| c 37 | 44.6 | 24.4 | 108 | 10 | US-09-894-711-20 | Sequence 20, Appl |
| 38 | 42.2 | 23.1 | 71 | 9 | US-09-861-687-15 | Sequence 15, Appl |
| 39 | 42.2 | 23.1 | 854 | 9 | US-09-954-531-989 | Sequence 989, App |
| 40 | 41.2 | 22.5 | 85 | 9 | US-09-861-687-14 | Sequence 14, Appl |
| 41 | 40.8 | 22.3 | 68 | 9 | US-09-861-687-16 | Sequence 16, Appl |
| 42 | 39.4 | 21.5 | 85 | 9 | US-09-861-687-25 | Sequence 25, Appl |
| 43 | 39 | 21.3 | 480 | 10 | US-09-921-398-38 | Sequence 38, Appl |
| 44 | 39 | 21.3 | 621 | 10 | US-09-921-398-40 | Sequence 40, Appl |
| c 45 | 37.6 | 20.5 | 112 | 9 | US-09-861-687-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-09-740-359-11
; Sequence 11, Application US/09740359
; Patent NO. US20010041787A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Borglum
; APPLICANT: Ludvigsen, Svend
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursor analogues having improved fermentation
; FILE REFERENCE: 6148.200-US
; CURRENT APPLICATION NUMBER: US/09740.359
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01869
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 550
; TYPE: DNA
; ORGANISM: TA57 leader fused with N-terminally extended....
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(486)
; OTHER INFORMATION: TA57 leader fused with N-terminally extended
; OTHER INFORMATION: AspB28IP(AspGlyLys)
US-09-740-359-11

Query Match 46.8% Score 85.6; DB 10; Length 550;
Best Local Similarity 72.8%; Pred. No. 6.3e-20;
Matches 131; Conservative 0; Mismatches 34; Indels 15; Gaps 1;
Qy 1 ATGTTTCGTTAATACGACACTGTGGCGTCTCACCTGGTAGAGCTCTGTACTGTTGTC 60
Db 325 AAGTTTCGTTAATACGACACTGTGGCGTCTCACCTGGTAGAGCTCTGTACTGTTGTC 384

[illegible]

ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4341.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:
LENGTH: 594 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

NAME/KEY: CDS
LOCATION: 109..522
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-861-687-18

Query Match 45.1%; Score 82.6; DB 9; Length 594;
Best Local Similarity 72.3%; Pred. No. 6.9e-19;
Matches 128; Conservative 0; Mismatches 34; Indels 15; Gaps 1;

Qy 4 TTCGTTAATCAGCACCTGTGGCGTCTCACCTGGTAGAAGCTCTGTACTGTTTGGCGT 63
Db 364 TTCGTTAACCAACACTTGTGGGTCCACCTGGTTGAAGCTTGTACTTGGTTGGCGT 423
Qy 64 GAACGTGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATCGTT 123
Db 424 GAAAGAGGTTTTCTTCTACACTCCCTAAGGCTGCTAAGGT-----ATTGTC 468
Qy 124 GAACAATCCTGCTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAAC TAG 180
Db 469 GAACAATCCTGCTACTCCATCTGCTCTGTGTACCAATTGGAAACTACTGCAACTAG 525

RESULT 5
US-09-740-359-9
; Sequence 9, Application US/09740359
; Patent No. US20010041787A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Borglum
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursor analogues having improved fermentation
; FILE REFERENCE: 6148.200-US
; CURRENT APPLICATION NUMBER: US/09/740,359
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01869
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(489)
; OTHER INFORMATION: Synthetic

US-09-740-359-9
Query Match 45.0%; Score 82.4; DB 10; Length 550;
Best Local Similarity 71.7%; Pred. No. 7.8e-19;
Matches 129; Conservative 0; Mismatches 36; Indels 15; Gaps 1;
Qy 1 ATGTTCTGTTAATCAGCACCTGTGGCGTCTCACCTGGTAGAAGCTCTGTACTGTTTGC 60
Db 328 AAGTTCGTTAACCAACACTTGTGGTCCACCTGGTTGAAGCTTTGTACTTGGTTTGT 387
Qy 61 GGTGAACGTGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
Db 388 GGTGAAGAGGTTTTCTTCTACACTGACAAGGACGGTAAGGT-----ATC 432
Qy 121 GTTGAACAATCCTGCTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAAC TAG 180
Db 433 GTTGAACAATCCTGCTACTTCTATCTGTTTGTGTACCAATTGGAAACTACTGTAAC TAG 492
RESULT 7
US-09-736-611-11
; Sequence 11, Application US/09736611
; Patent No. US20010023069A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas
; APPLICANT: Ludvigsen, Svend

US-09-740-359-9
Query Match 45.0%; Score 82.4; DB 10; Length 550;
Best Local Similarity 71.7%; Pred. No. 7.8e-19;
Matches 129; Conservative 0; Mismatches 36; Indels 15; Gaps 1;

Qy 1 ATGTTCTGTTAATCAGCACCTGTGGCGTCTCACCTGGTAGAAGCTCTGTACTGTTTGC 60
Db 328 AAGTTCGTTAACCAACACTTGTGGTCCACCTGGTTGAAGCTTTGTACTTGGTTTGT 387
Qy 61 GGTGAACGTGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
Db 388 GGTGAAGAGGTTTTCTTCTACACTGACAAGGACGGTAAGGT-----ATC 432
Qy 121 GTTGAACAATCCTGCTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAAC TAG 180
Db 433 GTTGAACAATCCTGCTACTTCTATCTGTTTGTGTACCAATTGGAAACTACTGTAAC TAG 492

RESULT 6
US-09-894-711-9
; Sequence 9, Application US/09894711
; Patent No. US2002037144A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas Borglum
; APPLICANT: Ludvigsen, Svend
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursor analogues having improved fermentation
; FILE REFERENCE: 6148.400-US
; CURRENT APPLICATION NUMBER: US/09/894,711
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PA 2000 00443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: PA 1999 01869
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/211,081
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: 60/181,450
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 09/740,359
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(489)
; OTHER INFORMATION: Synthetic

US-09-894-711-9
Query Match 45.0%; Score 82.4; DB 10; Length 550;
Best Local Similarity 71.7%; Pred. No. 7.8e-19;
Matches 129; Conservative 0; Mismatches 36; Indels 15; Gaps 1;
Qy 1 ATGTTCTGTTAATCAGCACCTGTGGCGTCTCACCTGGTAGAAGCTCTGTACTGTTTGC 60
Db 328 AAGTTCGTTAACCAACACTTGTGGTCCACCTGGTTGAAGCTTTGTACTTGGTTTGT 387
Qy 61 GGTGAACGTGTTTTTCTACACCCCGAAACCGGTGGTCCGGGTAAACGTGGCATC 120
Db 388 GGTGAAGAGGTTTTCTTCTACACTGACAAGGACGGTAAGGT-----ATC 432
Qy 121 GTTGAACAATCCTGCTACTAGCATCTGCTCTCTACACAGCTGGAGAACTATTGTAAC TAG 180
Db 433 GTTGAACAATCCTGCTACTTCTATCTGTTTGTGTACCAATTGGAAACTACTGTAAC TAG 492
RESULT 7
US-09-736-611-11
; Sequence 11, Application US/09736611
; Patent No. US20010023069A1
; GENERAL INFORMATION:
; APPLICANT: Kjeldsen, Thomas
; APPLICANT: Ludvigsen, Svend


```

; TYPE: DNA
; ORGANISM: N-terminal extension
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(489)
US-09-736-611-9

Query Match      42.4%; Score 77.6; DB 10; Length 550;
Best Local Similarity 70.0%; Pred. No. 3 4e-17;
Matches 126; Conservative 0; Mismatches 39; Indels 15; Gaps

Qy 1 ATGTTTCGTTAATCAGCACCTGTGGCGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTGTC 60
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 328 AAGTTCGTTAACCAACACTTGTGTGGTGTCCACTTGGTTGAAGCTTTGTACTTGGTTTGT 387

Qy 61 GGTGAACGCTGGTTTTTCTACACCCCGAACC GGTTGGTCCGGTAAACGTGGCAGTC 120
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 388 GGTGAAGAGAGTTTCTTCTACACTGACAAAGAAATGGAAGGGT-----ATC 432

Qy 121 GTTGAAACAATGCTCTACTAGCATCTGCTCTCTCTACACGCTGAGAACTATTCTGAAC TAG 180
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 433 GTTGAACAATGTTGACTTCTTCTCTTGTCTTTGTACCAATTGGAAAACTACTGTAACTAG 492

RESULT 12
US-09-736-611-5
; Sequence 5, Application US/09736611
; Patent No. US20010023069A1
; GENERAL INFORMATION:
; APPLICANT: Kjelsgen, Thomas
; APPLICANT: Ludvigsen, Svend
; APPLICANT: Kaarsholm, Niels
; TITLE OF INVENTION: Method For Making Insulin Precursors and
; TITLE OF INVENTION: Insulin Precursor Analogs
; FILE REFERENCE: 6058.200-US
; CURRENT APPLICATION NUMBER: US/09/736,611
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/181,443
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/211,441
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: PA 1999 01868
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: PA 2000 00440
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 600
; TYPE: DNA
; ORGANISM: N-terminally extension
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (114)...(545)
US-09-736-611-5

Query Match      39.6%; Score 72.4; DB 10; Length 600;
Best Local Similarity 69.4%; Pred. No. 2.1e-15;
Matches 125; Conservative 0; Mismatches 31; Indels 24; Gaps

Qy 1 ATGTTTCGTTAATCAGCACCTGTGGCGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGC 60
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 393 AAGTTCGTTAACCAACACTTGTGTGGTGTCTACTTGGTTGAAGCTTTGTACTTGGTTTGT 452

Qy 61 GGTGAACGCTGGTTTTTCTACACCCCGAACC GGTTGGTCCGGTAAACGTGGCAGTC 120
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 453 GGTGAAGAGAGTTTCTTCTACACTGACAAG-----GGTATC 488

Qy 121 GTTGAAACAATGCTGTACTAGCATCTGCTCTCTCTCTACACGCTGAGAACTATTCTGAAC TAG 180
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 489 GTTGAACAATGTTGACTTCTTCTCTTGTCTTTGTACCAATTGGAAAACTACTGTAACTAG 548

```

Job time : 69 secs

THIS PAGE BLANK (13770)

Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. J. Ferrer In vivo mass-excised to pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
Trace considered overall poor quality
Seq primer: -400P from Gibco
High quality sequence stop: 1.

FEATURES
SOURCE

Location/Qualifiers
1. .328

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5636897"
/clone_lib="Human insulinaoma"
/tissue_type="insulinaoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: pancreas; Vector: p Bluescript SK-; Site: 1:
xhoI; Site: 2: EcoRI; Constructed with lambda ZAPII system
(Stratagene) by Dr. J. Ferrer. in vivo mass-excised to
p Bluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."
61 a 106 c 101 q 60 t

```

[illegible]

61 a 106 c 101 g 60 t

Query Watch 37.28; Score 68; DB 13; Length 328;
Best Local Similarity 76.9%; Pred. No. 1.2e-11;
Matches 83; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

[illegible]

| | | | |
|----|-----|--|-----|
| Qy | 64 | GAACGTGGTTTTTTTCTACACCCCGAAAAACCGTGGTGGTCCGGGTAAA | 111 |
| | | | |
| Db | 196 | GAACGAGGGCTTTTTTCTACACCCCAAGACGCGCCCGGAGGCGAGAGAAA | 149 |

RESULT 2
AW583939/C

| | |
|------------|---|
| RESULT 2 | |
| AW583939/c | |
| LOCUS | |
| DEFINITION | 440 bp mRNA linear EST 13-MAR-2002 |
| | la07b07.x1 Human Pancreatic Islets Homo sapiens cDNA clone |
| | IMAGE:5637589 3' similar to gb:X70508 INSULIN PRECURSOR (HUMAN);, |
| | mRNA sequence |

ACCESSION
VERSION

Accession
Version
Keywords
Source

REFERENCE
AUTHORS

| REFERENCE | AUTHORS | TITLE |
|--------------------|--|------------------------------------|
| 1 (Pages 1 to 440) | Mellon, D., Martin, J., Clifton, S., Hillier, L., Marra, M., Page, D., Wylie, T., West, J., Blaustein, A., Schmitt, A., Theisinger, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y., and Bowers, Y. | WashU-Harvard Pancreas EST Project |
| | | Unpublished (2000) |

**JOURNAL
COMMENT**

COMMENT
published (2000)
Other_ESTS: ia07b07.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812
Fax: 617-495-8557

Fax: 617-495-8337
Email: dmelton@biohp.harvard.edu
Libraries were constructed by Dr. Douglas Melton

DNA sequencing by: Washington University Genome Sequencing Center
For information on obtaining a clone please contact: Juliana Brown
(brownjefas.harvard.edu)
This sequence now available from the IMAGE consortium, for clone
orders contact: info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 118.

[illegible]

Location/Qualifiers
1. .440

```
1. 340
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5637589"
/clone_lib="Human Pancreatic Islets"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="NH10B"
```

```

/notes=Organs: Pancreas; Vector: pSPORT1; Site_1: Not 1;
Site_2: Sal 1; Library constructed using SuperScript
P12; Library kit (Life Technologies). cDNA made by
oligo-dT priming. Size-selected by column fractionation;
average insert size 1.08 kb. Primary library,
unamplified."

```

| | | | | |
|------------|--------------|-------|-------|-------|
| BASE COUNT | 77 a | 129 c | 134 g | 100 t |
| ORIGIN | unpublished. | | | |

| | | | | |
|--------------------------|-------|--------------------|-----------|-------------|
| Query Match | 36.9% | Score 67.6; | DB 10; | Length 440; |
| Best Local Similarity | 77.4% | Pred. No. 1.8e-11; | | |
| Matches 82; Conservative | 0; | Mismatches 24; | Indels 0; | Gaps 0; |

QY 4 TTCGTTAATCAGCACCTGTGCGGCTCTCACCTGGTAGAAGCTCTGTACCTGGTTTGCGET 63

Db

353 TTTGTGAACCAACACCTGTGCGGTTACACACCTGGTGGAAAGTCTCTACTAGTGTGCGGG

64 GAACGTGGTTTTTCTACACCCCGAAAACCGGTGGTGGTCCGGGGA 109

Db 293 GAACGAGGCTTCTTCTACACACCCAGACCCGCTGGGAGGCCAGAGGA 248

RESULT 3
C06608/C

| RESULT 3 | LOCUS | 337 bp | mRNA | linear | EST 23-AUG-1996 |
|------------|--------|---------------------------|-------------------|---|-----------------|
| C06608/c | C06608 | Rat pancreatic islet cDNA | Rattus norvegicus | cDNA similar to insulin 1, mRNA sequence. | |
| DEFINITION | | | | | |

ACCESSION
VERSION

| ACCESSION | VERSION | INSTRUMENT | SEQUENCE |
|-----------|----------|------------|----------|
| C06608 | C06608.1 | GT-1503384 | |

VERSION
KEYWORDS
SOURCE

| VERSION | KEYWORDS | SOURCE | ORGANISM |
|-----------|------------|----------------------------------|---|
| C000061.1 | GI:1505364 | Norway rat. Rattus norvegicus | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. |

REFERENCE
AUTHORS

| | |
|-----------|---|
| REFERENCE | 1 (bases 1 to 337) |
| AUTHORS | Takeda,J. |
| TITLE | Large scale collection of expressed sequence tags (ESTs) from rat pancreatic islet cDNA library |
| JOURNAL | Unpublished (1996) |
| COMMENT | Contact : Jun Takeda |

JOURNAL
COMMENT

JOURNAL
COMMENT

Unpublished (1996)
Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@b.gunma-u.ac.jp.

FEATURES

```

FEATURES
  source
    1. .337
      Location/Qualifiers
        /organism="Rattus norvegicus"
        /db_xref="taxon:10116"
        /clone_lib="Rat pancreatic islet cDNA"
        /tissue_type="pancreatic islet"

```

BASE COUNT

| CIE ECORI SITE | | 4 others | |
|----------------|------|----------|------|
| 74 a | 97 c | 94 g | 68 t |
| 1 | 1 | 1 | 1 |
| 2 | 2 | 2 | 2 |
| 3 | 3 | 3 | 3 |
| 4 | 4 | 4 | 4 |
| 5 | 5 | 5 | 5 |
| 6 | 6 | 6 | 6 |
| 7 | 7 | 7 | 7 |
| 8 | 8 | 8 | 8 |
| 9 | 9 | 9 | 9 |
| 10 | 10 | 10 | 10 |
| 11 | 11 | 11 | 11 |
| 12 | 12 | 12 | 12 |
| 13 | 13 | 13 | 13 |
| 14 | 14 | 14 | 14 |
| 15 | 15 | 15 | 15 |
| 16 | 16 | 16 | 16 |
| 17 | 17 | 17 | 17 |
| 18 | 18 | 18 | 18 |
| 19 | 19 | 19 | 19 |
| 20 | 20 | 20 | 20 |
| 21 | 21 | 21 | 21 |
| 22 | 22 | 22 | 22 |
| 23 | 23 | 23 | 23 |
| 24 | 24 | 24 | 24 |
| 25 | 25 | 25 | 25 |
| 26 | 26 | 26 | 26 |
| 27 | 27 | 27 | 27 |
| 28 | 28 | 28 | 28 |
| 29 | 29 | 29 | 29 |
| 30 | 30 | 30 | 30 |
| 31 | 31 | 31 | 31 |
| 32 | 32 | 32 | 32 |
| 33 | 33 | 33 | 33 |
| 34 | 34 | 34 | 34 |
| 35 | 35 | 35 | 35 |
| 36 | 36 | 36 | 36 |
| 37 | 37 | 37 | 37 |
| 38 | 38 | 38 | 38 |
| 39 | 39 | 39 | 39 |
| 40 | 40 | 40 | 40 |
| 41 | 41 | 41 | 41 |
| 42 | 42 | 42 | 42 |
| 43 | 43 | 43 | 43 |
| 44 | 44 | 44 | 44 |
| 45 | 45 | 45 | 45 |
| 46 | 46 | 46 | 46 |
| 47 | 47 | 47 | 47 |
| 48 | 48 | 48 | 48 |
| 49 | 49 | 49 | 49 |
| 50 | 50 | 50 | 50 |
| 51 | 51 | 51 | 51 |
| 52 | 52 | 52 | 52 |
| 53 | 53 | 53 | 53 |
| 54 | 54 | 54 | 54 |
| 55 | 55 | 55 | 55 |
| 56 | 56 | 56 | 56 |
| 57 | 57 | 57 | 57 |
| 58 | 58 | 58 | 58 |
| 59 | 59 | 59 | 59 |
| 60 | 60 | 60 | 60 |
| 61 | 61 | 61 | 61 |
| 62 | 62 | 62 | 62 |
| 63 | 63 | 63 | 63 |
| 64 | 64 | 64 | 64 |
| 65 | 65 | 65 | 65 |
| 66 | 66 | 66 | 66 |
| 67 | 67 | 67 | 67 |
| 68 | 68 | 68 | 68 |
| 69 | 69 | 69 | 69 |
| 70 | 70 | 70 | 70 |
| 71 | 71 | 71 | 71 |
| 72 | 72 | 72 | 72 |
| 73 | 73 | 73 | 73 |
| 74 | 74 | 74 | 74 |
| 75 | 75 | 75 | 75 |
| 76 | 76 | 76 | 76 |
| 77 | 77 | 77 | 77 |
| 78 | 78 | 78 | 78 |
| 79 | 79 | 79 | 79 |
| 80 | 80 | 80 | 80 |
| 81 | 81 | 81 | 81 |
| 82 | 82 | 82 | 82 |
| 83 | 83 | 83 | 83 |
| 84 | 84 | 84 | 84 |
| 85 | 85 | 85 | 85 |
| 86 | 86 | 86 | 86 |
| 87 | 87 | 87 | 87 |
| 88 | 88 | 88 | 88 |
| 89 | 89 | 89 | 89 |
| 90 | 90 | 90 | 90 |
| 91 | 91 | 91 | 91 |
| 92 | 92 | 92 | 92 |
| 93 | 93 | 93 | 93 |
| 94 | 94 | 94 | 94 |
| 95 | 95 | 95 | 95 |
| 96 | 96 | 96 | 96 |
| 97 | 97 | 97 | 97 |
| 98 | 98 | 98 | 98 |
| 99 | 99 | 99 | 99 |
| 100 | 100 | 100 | 100 |

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 4, 2003, 07:00:27 ; Search time 2216 Seconds
(without alignments)
1337.443 Million cell updates/sec

Title: US-09-706-690-3
Perfect score: 183
Sequence: 1 atgttcgtaatcagcacct.....agaactattgtaactagtaa 183

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pln.*
- 21: em_gss_vrt.*
- 22: em_gss_fun.*
- 23: em_gss_mam.*
- 24: em_gss_mus.*
- 25: em_gss_Other.*
- 26: em_gss_pro.*
- 27: em_gss_rtd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|-----------|-------|-------|--------|----|----------|--------------------|--|
| Result | Query | | | | | | |
| No. | Score | Match | Length | DB | ID | Description | |
| C 1 | 68 | 37.2 | 328 | 13 | BM512022 | BM512022 i177f09.x | |
| C 2 | 67.6 | 36.9 | 440 | 10 | AW583939 | AW583939 ia07b07.x | |
| C 3 | 67.2 | 36.7 | 337 | 14 | C06608 | C06608 C06608 Rat | |
| C 4 | 67.2 | 36.7 | 397 | 14 | C08806 | C08806 C06806 Rat | |
| C 5 | 67.2 | 36.7 | 411 | 14 | BQ271766 | BQ271766 iK18b05.x | |
| C 6 | 67.2 | 36.7 | 415 | 14 | C06513 | C06513 C06513 Rat | |

| | | | | | | |
|------|------|------|-----|----|----------|--------------------|
| C 7 | 67.2 | 36.7 | 425 | 14 | C06514 | C06514 Rat |
| C 8 | 67.2 | 36.7 | 429 | 14 | C06642 | C06642 Rat |
| C 9 | 66.8 | 36.5 | 396 | 14 | C06974 | C06974 Rat |
| C 10 | 66.6 | 36.4 | 196 | 13 | BM508175 | BM508175 i140d11.x |
| C 11 | 66.6 | 36.4 | 328 | 13 | BM508175 | BM508175 i140d11.x |
| C 12 | 66.6 | 36.4 | 341 | 14 | C06750 | C06750 Rat |
| C 13 | 66.6 | 36.4 | 355 | 14 | BQ129971 | BQ129971 i192ell.x |
| C 14 | 66.6 | 36.4 | 359 | 14 | C06950 | C06950 Rat |
| C 15 | 66.6 | 36.4 | 418 | 13 | BM565516 | BM565516 ih26c01.x |
| C 16 | 66.6 | 36.4 | 418 | 14 | C06840 | C06840 Rat |
| C 17 | 66.6 | 36.4 | 421 | 14 | C07149 | C07149 Rat |
| C 18 | 66.6 | 36.4 | 423 | 10 | AW583178 | AW583178 ia11h05.y |
| C 19 | 66.6 | 36.4 | 425 | 13 | BM505743 | BM505743 ih12h11.y |
| C 20 | 66.6 | 36.4 | 427 | 14 | BQ270312 | BQ270312 iK09c05.y |
| C 21 | 66.6 | 36.4 | 428 | 14 | C06536 | C06536 Rat |
| C 22 | 66.6 | 36.4 | 430 | 14 | BQ549935 | BQ549935 i103d06.y |
| C 23 | 66.6 | 36.4 | 431 | 13 | BM567265 | BM567265 iD87e05.y |
| C 24 | 66.6 | 36.4 | 434 | 14 | C06831 | C06831 Rat |
| C 25 | 66.6 | 36.4 | 435 | 14 | BQ130182 | BQ130182 i192a01.y |
| C 26 | 66.6 | 36.4 | 436 | 14 | BQ549123 | BQ549123 iK89a07.y |
| C 27 | 66.6 | 36.4 | 447 | 13 | BM567267 | BM567267 iD87e08.y |
| C 28 | 66.6 | 36.4 | 452 | 14 | C06788 | C06788 Rat |
| C 29 | 66.6 | 36.4 | 453 | 14 | BQ581445 | BQ581445 i108g12.y |
| C 30 | 66.6 | 36.4 | 473 | 10 | AW583422 | AW583422 ia02e06.x |
| C 31 | 66.6 | 36.4 | 492 | 14 | BQ478368 | BQ478368 iK81a09.y |
| C 32 | 66.6 | 36.4 | 555 | 14 | C07148 | C07148 Rat |
| C 33 | 66.6 | 36.4 | 564 | 14 | C06914 | C06914 Rat |
| C 34 | 66.2 | 36.2 | 422 | 14 | BQ787802 | BQ787802 iM13d01.y |
| C 35 | 66.2 | 36.2 | 428 | 13 | BM509457 | BM509457 iH18f08.y |
| C 36 | 66 | 36.1 | 238 | 10 | AW583815 | AW583815 ia05a02.y |
| C 37 | 66 | 36.1 | 459 | 14 | BQ478188 | BQ478188 iK82b03.x |
| C 38 | 66 | 36.1 | 459 | 14 | BQ478332 | BQ478332 iK80e06.y |
| C 39 | 66 | 36.1 | 460 | 13 | BM512139 | BM512139 iJ75f09.y |
| C 40 | 65.8 | 36.0 | 326 | 13 | BM272549 | BM272549 iG99d07.x |
| C 41 | 65.8 | 36.0 | 351 | 14 | BQ477429 | BQ477429 iK86f10.x |
| C 42 | 65.8 | 36.0 | 403 | 14 | BQ614047 | BQ614047 i102h11.x |
| C 43 | 65.8 | 36.0 | 405 | 13 | BM565131 | BM565131 iJ73g11.y |
| C 44 | 65.8 | 36.0 | 412 | 14 | C06600 | C06600 Rat |
| C 45 | 65.8 | 36.0 | 434 | 13 | BM504395 | BM504395 ih22c07.y |

ALIGNMENTS

RESULT 1
BM512022/c
LOCUS i177f09.x1 Human Insulinoma Homo sapiens cDNA clone IMAGE:5636897
DEFINITION 3' similar to SW:INS_HUMAN P01308 INSULIN PRECURSOR. [1] ; , mRNA
sequence.
ACCESSION BM512022
VERSION BM512022.1 GI:18683165
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 328)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvillii,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Unpublished (2000)
TITLE Endocrine Pancreas Consortium
JOURNAL Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
COMMENT Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 415)
Query Match 36.7%; Score 67.2; DB 14; Length 415;
Best Local Similarity 80.4%; Pred. No. 2.4e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
JOURNAL
COMMENT
Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES
source
1..415
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="Rat pancreatic islet cDNA"
/tissue_type="pancreatic islet"
/notes="Vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XhoI;
mRNA was prepared from normal rat islets. cDNA was
directionally synthesized from the Xho I in the vector to
the EcoRI site"
BASE COUNT 82 a 130 c 111 g 90 t 2 others
ORIGIN
Query Match 36.7%; Score 67.2; DB 14; Length 415;
Best Local Similarity 80.4%; Pred. No. 2.4e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 4 TTCCTTAATCAGCACCCTGCTCCTCCTGCTAGAGCTCTGTACCTGTTGCGGT 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 100 TTGTCTAAACAGCACCCTTGTGGTCCCTCCTGCTGAGGCTCTGTACCTGTTGCGGT 159
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 GAACGTGTTTTTCTACACCCGAAACCGGTGGT 100
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 160 GAACGTGTTTTTCTACACACCCCAATCCCGTGTG 196
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 7
C06514/c
LOCUS
DEFINITION
C06514 Rat pancreatic islet cDNA Rattus norvegicus cDNA similar to
insulin 1, mRNA sequence.
ACCESSION
C06514.1 GI:1503290
VERSION
EST.
KEYWORDS
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 425)
AUTHORS
Takeda,J.
TITLE
Large scale collection of expressed sequence tags (ESTs) from rat
pancreatic islet cDNA library
JOURNAL
COMMENT
Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES
source
1..425
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="Rat pancreatic islet cDNA"
/tissue_type="pancreatic islet"
/notes="Vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XhoI;
mRNA was prepared from normal rat islets. cDNA was
directionally synthesized from the Xho I in the vector to
the EcoRI site"

```

```

BASE COUNT 101 a 113 c 122 g 81 t 8 others
ORIGIN
Query Match 36.7%; Score 67.2; DB 14; Length 425;
Best Local Similarity 80.4%; Pred. No. 2.4e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 4 TTCCTTAATCAGCACCCTGCTCCTCCTGCTAGAGCTCTGTACCTGTTGCGGT 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 TTGTCTAAACAGCACCCTTGTGGTCCCTCCTGCTGAGGCTCTGTACCTGTTGTTGGG 243
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 GAACGTGTTTTTCTACACCCGAAACCGGTGGT 100
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 GAACGTGTTTTTCTACACACCCCAAGTCCCGTNGTG 206
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 8
C06642/c
LOCUS
DEFINITION
C06642 Rat pancreatic islet cDNA Rattus norvegicus cDNA similar to
insulin 1, mRNA sequence.
ACCESSION
C06642
VERSION
C06642.1 GI:1503418
KEYWORDS
EST.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 429)
AUTHORS
Takeda,J.
TITLE
Large scale collection of expressed sequence tags (ESTs) from rat
pancreatic islet cDNA library
JOURNAL
COMMENT
Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Fax: 272-20-8896
Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES
source
1..429
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone_lib="Rat pancreatic islet cDNA"
/tissue_type="pancreatic islet"
/notes="Vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XhoI;
mRNA was prepared from normal rat islets. cDNA was
directionally synthesized from the Xho I in the vector to
the EcoRI site"
BASE COUNT 95 a 119 c 118 g 88 t 9 others
ORIGIN
Query Match 36.7%; Score 67.2; DB 14; Length 429;
Best Local Similarity 80.4%; Pred. No. 2.4e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 4 TTCCTTAATCAGCACCCTGCTCCTCCTGCTAGAGCTCTGTACCTGTTGCGGT 63
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 TTGTCTAAACAGCACCCTTGTGGTCCCTCCTGCTGAGGCTCTGTACCTGTTGTTGGG 268
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 64 GAACGTGTTTTTCTACACCCGAAACCGGTGGT 100
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 GAACGTGTTTTTCTACACACCCCAAGTCCCGTNGTG 231
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 9
C06974
LOCUS
DEFINITION
C06974 Rat pancreatic islet cDNA Rattus norvegicus cDNA similar to
insulin 1, mRNA sequence.
ACCESSION
C06974.1 GI:1503750
VERSION
C06974.1 GI:1503750

```


source

RÉSULT 13

BASE COUNT

Oxy

REFERENCE

1. (bases 1 to 355)

RESULT 12

ACCESSION

ORGANISM

| AUTHORS | TITLE |
|---------|-------|
| ... | ... |

FEATURES

university protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library.

TITLE Large scale collection of expressed sequence tags (ESTs) from rat pancreatic islet cDNA library
JOURNAL Unpublished (1996)
COMMENT Contact: Jun Takeda
Institute for Molecular and Cellular Regulation, Gunma University
3-39-15 Showa-machi, Maebashi Gunma 371, Japan
Tel: 272-20-8856
Tel: 272-20-8856
Fax: 272-20-8896
Email: ttakeda@sh.gunma-u.ac.jp

BASE COUNT
ORIGIN

36 a 60 c 66 g 34 t

xmol; Site_2: BCORI; Constructed with lambda ZAP11 system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
pBluescript SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library. "

| | | | | | | | | | | |
|-----------------------|-----|---|-----|------------|-----|-----------|----------|--------|----|-------------|
| | | 36 | a | 60 | c | 66 | g | 34 | t | |
| BASE COUNT: | | | | | | | | | | |
| ORIGIN | | | | | | | | | | |
| Query Match | | | | 36.4%; | | Score | 66.6; | DB 13; | | Length 196; |
| Best Local Similarity | | | | 80.4%; | | Pred. No. | 2.9e-11; | | | |
| Matches | 78; | Conservative | 0; | Mismatches | 19; | Indels | 0; | Gaps | 0; | |
| QY | 4 | TTGCGTTAATCAGCACCTGTGGCGCTCTACCTGGTGAAGCTCTGTACCTGGTTGCGET | 63 | | | | | | | |
| | | | | | | | | | | |
| Dd | 124 | TTTGTGAACCAACACCTGTGGCGCTCACACCTGGTGAAGCTCTCTACCTAGTGTGCGGG | 65 | | | | | | | |
| | | | | | | | | | | |
| QY | 64 | GACCTGGTGGTTTTTTCTACACCCCGAAACCGGTGGTG | 100 | | | | | | | |
| | | | | | | | | | | |
| Dd | 64 | GACACGAGGCTTTTTCCTACACACCAAGACCCGCCGG | 28 | | | | | | | |
| | | | | | | | | | | |

| | | | | | |
|------------|------------|-----------------------------------|-------------------------------|------------|-----------------|
| RESULT 11 | BI711284/c | BI711284 | 328 bp | linear | EST 11-MAR-2002 |
| LOCUS | | id96e03.x1 | Human insulinoma Homo sapiens | cDNA clone | IMAGE:5023588 |
| DEFINITION | | 3' similar to SW:INS_HUMAN P01308 | INSULIN PRECURSOR. [1] ; | mrna | |
| | | sequence. | | | |
| ACCESSION | | BI711284 | | | |
| VERSION | | BI711284.1 | GI:15686979 | | |
| KEYWORDS | | EST. | | | |
| SOURCE | | human. | | | |

| | |
|-----------|---|
| ORGANISM | homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 328) |
| AUTHORS | Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,X. and Bowers,Y. |
| TITLE | Endocrine Pancreas Consortium |
| JOURNAL | Unpublished (2000) |

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biochp.harvard.edu
 Library was constructed by Dr. J. Ferrer In vivo mass-excised to

```

BASE COUNT      63 a  115 c  114 g  63 t
ORIGIN
Query Match      36.4%; Score 66.6; DB 14; Length 355;
Best Local Similarity 80.4%; Pred. No. 3.6e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTGGAGAGCTCTGTACCTGGTTTGGCGGT 63
    || || || || || || || || || || || || || || || || || || || || || ||
Db 279 TTGTGGAACCAACACCTGTGGGCTCACACCTGGTGGAGAGCTCTGTACCTAGTGTGGGG 220
    || || || || || || || || || || || || || || || || || || || || || ||

Qy 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGT 100
    || || || || || || || || || || || || || || || || || || || || || ||
Db 219 GAACGAGGCTTTTCTACACACCCCAAGACCCGCCGG 183
    || || || || || || || || || || || || || || || || || || || || || ||

RESULT 14
C06950 LOCUS      C06950 Rat pancreatic islet cDNA Rattus norvegicus EST 23-AUG-1996
DEFINITION insulin 1, mRNA sequence.
ACCESSION C06950
VERSION C06950.1 GI:1503726
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 359)
AUTHORS Takeda, J.
TITLE Large scale collection of expressed sequence tags (ESTs) from rat
        pancreatic islet cDNA library
JOURNAL Unpublished (1996)
COMMENT Contact: Jun Takeda
        Institute for Molecular and Cellular Regulation, Gunma University
        3-39-15 Shoua-machi, Maebashi Gunma 371, Japan
        Tel: 272-20-8856
        Fax: 272-20-8896
        Email: jtakeda@sb.gunma-u.ac.jp.
FEATURES
        Location/Qualifiers
            1..359
                /organism="Rattus norvegicus"
                /db_xref="taxon:10116"
                /clone_lib="Rat pancreatic islet cDNA"
                /tissue_type="pancreatic islet"
                /note="Vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XhoI;
                mRNA was prepared from normal rat islets. cDNA was
                directionally synthesized from the Xho I in the vector to
                the EcoRI site"
BASE COUNT      70 a  113 c  99 g  77 t
ORIGIN
Query Match      36.4%; Score 66.6; DB 14; Length 359;
Best Local Similarity 80.4%; Pred. No. 3.6e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTGGAGAGCTCTGTACCTGGTTTGGCGGT 63
    || || || || || || || || || || || || || || || || || || || || || ||
Db 49 TTGTGCAACACGACCTTTGTGGTCTCACCTGGTGGAGGCTCTGTACCTGGTGTGGGG 108
    || || || || || || || || || || || || || || || || || || || || || ||

Qy 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGT 100
    || || || || || || || || || || || || || || || || || || || || || ||
Db 109 GAACGTGGTTTTTCTACACACCCCAAGTCCCGTCGTG 145
    || || || || || || || || || || || || || || || || || || || || || ||

RESULT 15
BM565516 LOCUS      BM565516 418 bp mRNA linear EST 20-FEB-2002
DEFINITION h26c01.x1 Human insulinoma Homo sapiens cDNA 3' similar to
          SW:INS_HUMAN P01308 INSULIN PROMOTOR. [1]; mRNA sequence.
ACCESSION BM565516
VERSION BM565516.1 GI:18825753

```

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 418)
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
          Lenishka, I., Scarce, M., Brestelli, J., Gradwohl, J., Clifton, S.,
          Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
          Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
          , M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T.
          , Jackson, Y. and Bowers, Y.
          Endocrine Pancreas Consortium
          Unpublished (2000)
          Other_ESTs: ih26c01.y1
          Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
          Endocrine Pancreas Consortium
          Harvard University, Howard Hughes Medical Institute
          Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
          MA 02138
          Tel: 617-495-1812
          Fax: 617-495-8557
          Email: dmelton@biohp.harvard.edu
          Library was constructed by Dr. J. Ferrer In vivo mass-excised to
          pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
          University Genome Sequencing Center For information on obtaining a
          clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
          Possible reversed clone: similarity on wrong strand
          Seq primer: -400p from Glibco.
          Location/Qualifiers
            1..418
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_lib="Human insulinoma"
                /tissue_type="insulinoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                pBluescript SK- by Dr. H. Inoue following the Washington
                University protocol
                (http://genome.wustl.edu/est/lambda_protocol.shtml).
                Please contact Hiroshi Inoue, MD/PhD for further
                information on this library (Metabolism Division, Permutt
                Laboratory, Washington University School of Medicine, Box
                8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
                is a Washington University Pancreas EST project library."
BASE COUNT      76 a  142 c  129 g  70 t  1 others
ORIGIN
Query Match      36.4%; Score 66.6; DB 13; Length 418;
Best Local Similarity 80.4%; Pred. No. 3.8e-11;
Matches 78; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 4 TTCGTTAATCAGCACCTGTGGGCTCTCACCTGGTGGAGAGCTCTGTACCTGGTTTGGCGGT 63
    || || || || || || || || || || || || || || || || || || || || || ||
Db 109 TTGTGGAACCAACACCTGTGGGCTCACACCTGGTGGAGAGCTCTGTACCTAGTGTGGGG 168
    || || || || || || || || || || || || || || || || || || || || || ||

Qy 64 GAACGTGGTTTTTCTACACCCCGAAACCGGTGGT 100
    || || || || || || || || || || || || || || || || || || || || || ||
Db 169 GAACGAGGCTTTTCTACACACCCCAAGACCCGCCGG 205
    || || || || || || || || || || || || || || || || || || || || || ||

Search completed: February 4, 2003, 09:28:03
Job time : 2229 secs

```

THIS PAGE BLANK (USPTO)